

SEQUENCE LISTING

<110> University of Iowa Research Foundation et al.

5

<120> Novel polypeptides and methods of their use

<130> 875.006WO1

10

<150> US 60/105,575

<151> 1998-10-26

<160> 65

15

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 2265

<212> DNA

20

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

25

<400> 1

| | | |
|----|--|-----|
| | gagggtagag aaatcgaatg ccccttgcac caaggtcggg ttgacgtttg cacaggcaaa | 60 |
| | gccctgtgag caccctgtgac acagaacatc aaaacatata cagtcaagat tgagaacctg | 120 |
| | cgcgtaataa ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt | 180 |
| | aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt | 240 |
| 30 | aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca | 300 |
| | tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc | 360 |
| | tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga | 420 |
| | gaacgacggt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggg | 480 |
| | gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg | 540 |
| 35 | ctccaacggt gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa | 600 |
| | taaaaaatgt ctgggggttg aagaagtcgc tcgcgtggag agcttccatg gcttcatcta | 660 |
| | cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg | 720 |
| | cctggaacct atgttcaagc attccggcgg tttagaactg gtcggctctc caggcaaggt | 780 |
| | tgtgatcaag gccaaactga aggcacccgc ggaaaacttt gtgggagatg cataccacgt | 840 |
| 40 | gggttgagc cagcgtctt cgttcgctc gggggagtct atcttctcgt cgctcgtcgt | 900 |
| | caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg | 960 |

```

catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020
gatggcattc ggaggcgcaa agcaggaaaag gctgaacaaa gaaattggcg atgttcgcgc 1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
ctcgggtggt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacctg 1200
5 cgccattgtc gaaaaagaca tgccctgagga tctcaagcgc cgcttggccg actctgttca 1260
gcgaacggtc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380
cggtgaggac gtatacggcg acgcgggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440
cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg 1500
10 ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggg ttccgcccac 1620
gacgcogaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
actacgctgc tgaccagga agcgcathtt ttggacattc aggttaccg tgcttggtta 1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
15 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100
20 cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg 2265

```

<210> 2

25 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

30 <223> A polypeptide encoded by SEQ ID NO:1

<400> 2

```

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
  1             5             10             15
35 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
      20             25             30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
      35             40             45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
40      50             55             60

```

Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
 65 70 75 80
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
 85 90 95
 5 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
 100 105 110
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
 115 120 125
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 10 130 135 140
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
 145 150 155 160
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
 165 170 175
 15 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
 180 185 190
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
 195 200 205
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
 20 210 215 220
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
 225 230 235 240
 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
 245 250 255
 25 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 30 290 295 300
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 35 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 355 360 365
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 40 370 375 380

10

<211> 9841

<212> DNA

<220>

<223> A modified DNA sequence

| | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
| | gaattcatca | ggaagacatt | caaatgaacg | taaacacataa | gggcagcgctc | tgtattttgcg | 60 |
| | gcagcgaaat | gctccctaaa | ttcctcattt | accccatctg | aggattgctt | tatgacagta | 120 |
| | aagtggattg | aagcagtcgc | tctttctgac | atccttgaag | gtgacgtcct | cggcgtgact | 180 |
| | gtcgaaggca | aggagctggc | gctgtatgaa | gttgaaggcg | aaatctacgc | taccgacaac | 240 |
| 25 | ctgtgcacgc | atggtttccgc | ccgcatgagt | gatggttatc | tcgagggtag | agaaatcgaa | 300 |
| | tgccccttgc | atcaaagtcg | gtttgacgtt | tgcacaggca | aagccctgtg | cgcacccgtg | 360 |
| | acacagaaca | tcaaaacata | tccagtcaag | attgagaacc | tgcgcgtaat | gattgatttg | 420 |
| | agctaagaat | tttaacagga | ggcaccgccg | gccttagagc | gtaatcacc | ccattccatc | 480 |
| | ttttttaggt | gaaaacatga | attacaataa | taaaatcttg | gtaagtgaat | ctgggtctgag | 540 |
| 30 | ccaaaagcac | ctgattcatg | gcgatgaaga | actttttcaa | catgaactga | aaaccatttt | 600 |
| | tgcgcggaac | tggctttttc | tcactcatga | tagcctgatt | cctgcccccg | gcgactatgt | 660 |
| | taccgcaaaa | atgggggattg | acgaggtcat | cgtctcccgg | cagaacgacg | gttcgattcg | 720 |
| | tgctttttctg | aacgtttgcc | ggcatcgtgg | aagacgcgtg | gtgagcgtgg | aagccggcaa | 780 |
| | tgccaaaggt | tttgtttgca | gctatcacgg | ctgggggcttc | ggctccaacg | gtgaactgca | 840 |
| 35 | gagcgttcca | tttgaaaaag | atctgtacgg | cgagtcgctc | aataaaaaat | gtctgggggtt | 900 |
| | gaaagaagtc | gctcgcgtgg | agagcttcca | tggcttcatc | tacggttgct | tcgaccagga | 960 |
| | ggcccctcct | cttatggact | atctgggtga | cgtcgtctgg | tacctggaac | ctatgttcaa | 1020 |
| | gcattccggc | ggtttagaac | tggtcggtcc | tccaggcaag | gttgtgatca | aggccaactg | 1080 |
| | gaaggcaccc | gcggaaaact | ttgtggggaga | tgcataccac | gtgggttgga | cgcacgcgctc | 1140 |
| 40 | tttcqcttcgc | tcqggggagt | ctatctttctc | gtcgcctcgt | ggcaatgcgg | cgctaccacc | 1200 |

| | | | | | | |
|--------------|------------|------------|------------|------------|-------------|------|
| tgaaggcgca | ggcttgcaaa | tgacctcaa | atacggcagc | ggcatgggtg | tggtgtggga | 1260 |
| cggatattca | ggtgtgcata | gcgagactt | ggttcggaa | ttgatggcat | tcggaggcgc | 1320 |
| aaagcaggaa | aggctgaaca | aagaaattgg | cgatgttcgc | gctcggattt | atcgagacca | 1380 |
| cctcaactgc | accgttttcc | cgaacaacag | catgctgacc | tgctcgggtg | ttttcaaagt | 1440 |
| 5atggaacccg | atcgacgcaa | acaccaccga | ggtctggacc | tacgccattg | tcgaaaaaga | 1500 |
| catgcctgag | gatctcaagc | gccgcttggc | cgactctgtt | cagcgaacgg | tcgggcctgc | 1560 |
| tggtctctgg | gaaagcgacg | acaatgacaa | tatggaaaca | gcttcgcaaa | acggcaagaa | 1620 |
| atatcaatca | agagatagtg | atctgctttc | aaaccttggg | ttcggtgagg | acgtatacgg | 1680 |
| cgacgcggtc | tatccaggcg | tcgtcggcaa | atcggcgatc | ggcgagacca | gttatcgtgg | 1740 |
| 10tttctaccgg | gcttaccagg | cacacgtcag | cagctccaac | tgggctgagt | tcgagcatgc | 1800 |
| ctctagtact | tggcatactg | aacttacgaa | gactactgat | cgctaacaga | cgagtcgacc | 1860 |
| atgatgatca | atattcaaga | agacaagctg | gtttccgccc | acgacgccga | agagattcct | 1920 |
| cgtttcttca | attgccacga | ctctgctttg | caacaagaag | ccactacgct | gctgaccacg | 1980 |
| gaagcgcat | tggtggacat | tcaggcttac | cgtgcttggg | tagagcactg | cgtgggggtca | 2040 |
| 15gaggtgcaat | atcaggtcat | ttcacgcgaa | ctgcgcgcag | cttcagagcg | tcgttataag | 2100 |
| ctcaatgaag | ccatgaacgt | ttacaacgaa | aattttcagc | aactgaaagt | tcgagttgag | 2160 |
| catcaactgg | atccgcaaaa | ctggggcaac | agcccgaagc | tgcgctttac | tcgctttatc | 2220 |
| accaacgtcc | aggccgcaat | ggacgtaa | gacaaagagc | tacttcacat | ccgctccaac | 2280 |
| gtcattctgc | accgggcacg | acgtggcaat | caggtcgatg | tcttctacgc | cgcccgggaa | 2340 |
| 20gataaatgga | aacgtggcga | aggtggagta | cgaaaattgg | tccagcgatt | cgtcgattac | 2400 |
| ccagagcgca | tacttcagac | gcacaatctg | atggtctttc | tgtgattcag | tgaccatttt | 2460 |
| tacaaatggt | cactgcaacc | gcggtcacca | ttaatcaaag | ggaatgtacg | tgtatgggca | 2520 |
| atcaacaagt | cgtttcgata | accggtgcag | gctcaggaat | cggtctcgaa | ctgggttcggg | 2580 |
| cctttaagtc | ggccggttat | tacgtatccg | ctctcgtacg | aaacgaggag | caagaggcgc | 2640 |
| 25ttctttgcaa | agagttcaag | gacgcactcg | agattgtagt | gggcgatgtc | cgggaccacg | 2700 |
| caacaaatga | gaagctgata | aagcaaaca | tcgatagatt | cggtcatctt | gattgtttta | 2760 |
| ttgcaaatgc | cggatctctg | gattacatgc | tgagcatcga | agagccttgg | gagaaaatat | 2820 |
| cgagcagttt | tgacgaaata | ttcgacatta | atgtcaagag | ctatttcagt | ggcatcagtg | 2880 |
| ccgccctgcc | ggaactgaaa | aagactaacg | gatcagtggt | gatgaccgct | tcggtgtcgt | 2940 |
| 30cccatgcggg | cgggtggtgg | ggttcttgct | acatcgccag | caagcatgcg | gtgctcggtg | 3000 |
| tggttaaggc | tttggcctac | gaattggccc | ccgaagtctg | cgtgaacgct | gtttcgccgg | 3060 |
| ggggcacccg | gacgtctctg | tgcggtcccg | cgagcgccgg | tttcgacaaa | atgcacatga | 3120 |
| aagacatgcc | cggcatcgac | gatatgatca | aaggtctcac | gcctcttggg | tttgagacca | 3180 |
| agcccgaaga | cgtggtggca | ccctatttgt | tgctggcttc | gcgaaagcaa | ggaaaattca | 3240 |
| 35tcaccggcac | cgtgattagc | attgatggcg | gtatggcgct | cggtcgcaag | tgagcttgta | 3300 |
| gccgatcaga | agttatagac | acatttcagg | tgacgcccc | tgaagacaaa | actgtttatc | 3360 |
| aataacgcct | ggatcgattc | tagtgaccag | cagaccttcg | agcgcataca | ccccgtcagc | 3420 |
| agcgatgtgg | tgactgagag | cgcaaacgcc | acagtgcagg | acgcgataaa | ggcggcgcaa | 3480 |
| gcggccgagg | aggcgttcaa | gacctggaag | gccgttgga | cttcagagcg | tcgccgcctt | 3540 |
| 40ctcctaaagg | tcgccgatgt | catggaaagt | aaaacacca | agttcatcga | agtgatggcc | 3600 |

| | | | | | | | |
|----|-------------|------------|-------------|-------------|------------|------------|------|
| | atggaggttg | gagcttcgc | cttttggcc | ggattcaacg | tccatgcgtc | tgccaatgtg | 3660 |
| | ttccgagagg | ctgcctcgtc | ggctacccaa | attcaggggtg | aaaccatccc | aacggacaaa | 3720 |
| | gccgaaacgc | tctcaatgac | actacgtcag | ccggtcggcc | cgatcctaag | catcgttcca | 3780 |
| | tggaacggca | ccgcagtgtc | tgcggcacga | gccatcgctt | atccgctggt | ctgtggcaac | 3840 |
| 5 | actgtggtgt | tcaaaggctc | tgaatttagt | cccgcgacgc | atgccctgat | cacccagtgc | 3900 |
| | gtgcaggaag | ccgggctgcc | cgctggcggtg | ctcaattacc | tcaactcttc | gcctgaccgt | 3960 |
| | tcgcccgaga | tcgctgacgc | actgatctct | gccaaggaga | tccgccgcat | caacttcacg | 4020 |
| | ggttcacccc | gcgtgggcag | cattatcgcg | cagaaagccg | cgcaaacctt | caagcgctgc | 4080 |
| | ctgctggagc | tcggcggcaa | gtccccgctt | attgttcttg | atgatgcaga | catcgatgcg | 4140 |
| 10 | cgggtcaagg | cagcgggtgt | cggtagcttc | ctgttccaag | gtcagatctg | catgtccact | 4200 |
| | gagcgcttga | tcgttgatga | gaagatagcc | gacgaatttg | tcgcaaaatt | tgtcgaaaaa | 4260 |
| | actaagcgct | tgagcgcagg | cgaccctgtc | gtaactggcg | actgcatcat | cggcccgatg | 4320 |
| | gtctcgccaa | attcgggtga | gcggatcaat | ggtttgttca | aagacgcgat | cgacaaaagg | 4380 |
| | gcaaaagttg | tttgcggcgg | cttggcccaa | ggtgcgctca | tgccggccac | gatcctggat | 4440 |
| 15 | cacgtcaaat | ctgacatgcg | gatttacgat | gaggagacct | ttggtcccat | caccgtggta | 4500 |
| | atccgttgta | aaggcgaagc | agaggccgtc | cgcattgcca | acgacagcgt | ctatggcctg | 4560 |
| | tcgtcgggcg | tatttggccg | cgacatcaac | cgcgctctac | gcgtgggtat | gtccatcgaa | 4620 |
| | tatggttctg | tacacatcaa | cggttcgacc | gtccagaacg | aggcgcaggc | tccttacgga | 4680 |
| | ggcaccaaga | acaccggcta | cgggcgcctt | gacggccgtg | ctgtaatcga | cgagttcaca | 4740 |
| 20 | gagatcaagt | ggctgaccat | cgaacctttc | gagcagcaat | atcccttctg | ataagcacta | 4800 |
| | actcccagga | atcaaactat | gagtaaagcaa | gctgcagtta | tcgagctcgg | atacatgggt | 4860 |
| | atctcgggtca | aggaccctga | tgctgggaaa | tcatttgcca | cggatatgct | aggtctgcaa | 4920 |
| | gttcttgatg | agggtgagaa | ggaccgtttc | tatctgcgga | tggattactg | gcatcatcgg | 4980 |
| | atcgtagtcc | atcacaacgg | acaggacgac | ttggagtacc | taggctggcg | tgtagecggc | 5040 |
| 25 | aagccggagt | tcgaagctct | gggtcaaaag | cttattgatg | ccggttacia | gatccgcctc | 5100 |
| | tgcgacaaa | ttgagggtca | ggagcgtatg | gtgttgggtc | tgatgaagac | agaagatccg | 5160 |
| | ggcggcaacc | cgaccgagat | attctggggc | cccgggatcg | acatgagcaa | cccgttccat | 5220 |
| | cccggtcgcc | ccctgcacgg | aaagtttgtg | accggtgacc | aaggcttggg | ccattgcctc | 5280 |
| | gttcgccaaa | ccgacgtcgc | agaagctcat | aagttttata | gcctgctggg | cttcctgtgg | 5340 |
| 30 | gacgtcgaat | accggattcc | gttgcccaac | ggcatgactg | ccgaactgtc | gttcatgcat | 5400 |
| | tgcaacgccc | gtgatcactc | cattgctttt | ggtgccatgc | ccgctgcaa | acgactcaat | 5460 |
| | cacttgatgc | ttgagtacac | ccatatggaa | gacttgggat | acacgcacca | acagtttgta | 5520 |
| | aagaacgaaa | ttgacattgc | cttgacgctt | ggcattcacg | ccaacgacaa | ggcgttgacg | 5580 |
| | ttctatgggtg | caacgccttc | gggctggctc | attgagcccg | gctggcgagg | tgccacggcc | 5640 |
| 35 | atagatgaag | cggagtatta | cgctcggcgac | atcttcggcc | atggcgtgga | ggccactgga | 5700 |
| | tatggcctgg | atgtaaaact | gagctaaaga | tgcgcgctcg | ttgggcgagg | ctctagtcca | 5760 |
| | gcatcttcat | acgcaaccaa | ccttgacggg | cgatgagatc | aaaggacgtt | aaagcgaagg | 5820 |
| | ggaagtgggt | cgggccatgc | gcataccgat | ccatgacatt | tgtttcatag | tatataggta | 5880 |
| | gataggtgaa | tcaagcgctt | agtcaactag | tggacacatc | tgttccatga | ggctatctac | 5940 |
| 40 | tatctattca | aaacaagaat | aataaatagg | atgaaaataa | taatgataaa | aagaacgatt | 6000 |

tgtcttgtgt atcctctatt ctgtttggca agccccacat gggccgaaga gtcgccttgg 6060
 acgtaccgta ttggtatgac taatgtagct ttcgatgcta gcgcaaaagt atacttaaatt 6120
 ggtcagcggg tgccaggagg aagcgtgat gcgagcgata acaacgcgct tacattcgac 6180
 ttcggctacg ccatcaacga ccagtggaat gtacgtgcga ttgtcggat tccgcctaca 6240
 5actaaagtga cgggcgcagg cacacttcct ggtatccagc tggggaaaat aacttacgct 6300
 ccaacagtat taacgttgaa ctataacctc cccgctttgg gtcccgttcg ccctcacata 6360
 ggtgcgggag tcaattacac gcggattttt gaaagtcggg acgctaattc aaaatcgttc 6420
 gatgccgacc acgcttggtc cccgcgcta catgttgggt ccgatattga cgttaaccgt 6480
 ggttggttcg ttagcattga tatccggaag ttatacctga aaaccgacgc atcagggtag 6540
 10ttggggccac aggaggctaa agcacgggta actcttgacc cactactaac ttcgatcgcg 6600
 atcggacgcc aattctgatg attctgttta aagttcttta tctatctaac cgcaaagggt 6660
 gtttccatgt cgaataaaat tatgaaaacg tcgcgtctta ccgccgaaga tatcaacggc 6720
 gcctggacta taatgccac accctcgacg cctgatgctt ctgattggcg cagcactgcc 6780
 accgtggact tagaagagac tgcccgcata gttgaagagc tgattgcagc tgggtgtcaac 6840
 15ggtattctaa gtatgggtac ttttgggtgag tgccgcacgt tgacctggga tgaaaaacgt 6900
 gattatgtct cgacgattgt cgagaccatt cgtggtcgcg tgccttattt ctgtggcacg 6960
 acagccttaa atacccgaga agtcatccgc cagacccgag agcttatcga tattggcgcc 7020
 aacggcacta tgctcggggg gccgatgtgg gtgaagatgg acctgcctac agcggttcag 7080
 ttctatcggt atgttcgaga tgccgtacca gaggtgccca ttgcgattta cgccaacccc 7140
 20gaagcattca agttcgactt ccctcgccca ttctgggcag agatgtccaa aattccgcag 7200
 gtagtgactg cgaagtatct aggcacgga atgcttgact tggacctgag actggcacc 7260
 aacatccgct tccttcccca cgaagatgac tattacgcgg ccgcacgcat caatcccgag 7320
 cgcataaacc cgttctggtc aagcggggcc atgtgcggcc cggctaccgc catcatgttg 7380
 cgtgacgaag tggtgccggc caagagcacc ggtgactggg ccaaggccaa agccatctcc 7440
 25gatgatatgc gtgcagccga ctcgacattg tttccgcgtg gcgacttttc ggagttctcg 7500
 aagtataata tcgggcttga aaaggcacgg atggatgcgg ctggttgggt caaggctggg 7560
 ccctgcgctc cgccctacaa ccttgttcca gaagactacc tcgctgggtc acagaaatca 7620
 ggcaaggctt gggccgcgct gcacgctaaa tacagtaatg aattgaagta gttcacctcc 7680
 gcagacctga gtgacagggt ggcgcagacg ctgaggggtg aggaattaag tgagctaaa 7740
 30cacatttctt gcgccaggca ttgccagatc agcaaagttt gctgatctgg cagtttcaaa 7800
 aatttggcg aaagctgata tcaggaatac gggataaagg cagtgcacca taacgacggg 7860
 gcgtgccatt cgtgatgaac gattttgcta ttgtgccgac ttctgttctt ggagtgttg 7920
 attgtgattg tcgattttta tttcgatttt ttgagtcctt tctcttactt ggccaacat 7980
 cgtttgtcaa agcttgcgca agactatggc ttttccattc gttattacgc aatcgatttg 8040
 35gcgcgagtta aaatagccat cggaaacgtt ggtccatcta atcgcgacct gatagtcaag 8100
 ctggactatt tgaaagtaga tttgcaacgg tgggcccagc tttacgaaat accgttggt 8160
 ttcccagcta actacaacag ccgacggatg aatactgggc tttattactc gggagccatg 8220
 gcacagactg gtgcctatgt gaatgtagta tttaatgcgg tttggggaga tggcatagct 8280
 ccagatttgg aaagcttgcc tgctctggta tctgaaaaac taggctggga tcgtagcgcc 8340
 40ttcgaggact ttatcagcag cgatgccgca acagagaggt atgacgagca gacacatgcc 8400

gcgatcgaac gcaaagtgtt cgggtgtgcca acgatgtttt tgggcatga aatgtggtg 8460
 ggaaacgacc gtctatttat gctcgagaac gcagtgggag gtgcgcctgt aaatggagaa 8520
 tagtcgctac ggagcgcttg tgcgggctaa atgccgatat aagtgggtga cctgatcggt 8580
 atttgctcga tacagcgctt tcaaaatcag cggctactga agtcagataa aaatgcggga 8640
 5ctacttcagg catcctgtgc gacacaaagt ttacctgta attgtccacc tattccgagt 8700
 ttggaatggg agctgactcg ctatgcgacc agcgatagcc taacaagaca tgcactactg 8760
 gtaacgggtg ggtgtgaagc tcctgcaaca atgtagcccc ttgatgtgtg tatttgctgc 8820
 gaggtgaagc acagatgctc ggagccgtac cggcttgtgg cgctaggctg gcaagtatga 8880
 gcaacgtaag tgggggttg ggcgcaatgg gaacaaaaa ccaacgcaag ccttaccagc 8940
 10gtcggttcggg gccttcctcc catgcctccg cctcgataaa gcagctgcgc atatcggtt 9000
 cctggctgat ctcggttagt aggtcatgca aggtcttgtc cagcgctcg tcgctccgat 9060
 acggaatggg cagctcgtaa tggccggtct ccgaccgctt catgccgtag ggctccaggc 9120
 agtagcgctc gatgttctcc gtggcccgt tccgaccgcg catgaacttg ctgttgttca 9180
 ccaccgccag gcgcagggtg acggtggcca cccgctcgac ggttgactct gccggtgacg 9240
 15cgatattgcg cttttgacct cgcgccaggg cgctcttctg gtacgtcccg atctcgacgc 9300
 cacggtggcg taggtagctg tacagggtgc tcttgagat gtgcaacttc tcgccgatgg 9360
 cgctgacgct caggcgcccc tcgcggtaca gggctctccg cgccatggcg gtggcctcgg 9420
 ccttggtggtg caggcccttg ggacggcgac cgatccggcc tcgagtcctg gccgccgaca 9480
 ggcccgctg agtccgctcg cggatcagct cgcgctcgaa ctgaacagg tgaacaccag 9540
 20gcgatcttg gcggtgggtg tgtcaatggg gtcgttcagg ctctgcaagc cgactttgcg 9600
 tgcagccagc tagccgacca actcaaccag gtgcttgagc gagcgaccga ggcgatccag 9660
 ctccagatc accacggcat cgcgcgctcg aacatgggct agcaacttgt ccaactccgg 9720
 ccgcgcgctt tttgcgccgc tggcgatgtc ttgatagatg cgttcgcacc cggcctgttt 9780
 cagggcacg acctggaggt cggcggttcta atcccgagtg ctcacccgcg tataaccgat 9840
 25c 9841

<210> 4

30 <211> 2515

<212> DNA

<213> Artificial Sequence

<220>

35 <223> A modified DNA sequence

<400> 4

gaattcatca ggaagacatt caaatgaacg taaacaataa gggcagcgtc tgtatttgcg 60
 gcagcgaaat gtccctaata ttccctcattt accccatctg aggattgctt tatgacagta 120
 40 aagtggattg aagcagtcgc tctttctgac atccttgaag gtgacgtcct cggcgtgact 180

| | | | | | | | |
|----|-------------|------------|------------|-------------|------------|-------------|------|
| | gtcgagggca | aggagctggc | gctgtatgaa | gttgaaggcg | aaatctacgc | taccgacaac | 240 |
| | ctgtgcacgc | atgggtccgc | ccgcatgagt | gatggttatc | tcgagggtag | agaaatcgaa | 300 |
| | tgccccctgc | atcaaggctg | gtttgacgtt | tgcacaggca | aagccctgtg | cgcacccgtg | 360 |
| | acacagaaca | tcaaaacata | tccagtcaag | attgagaacc | tgcgcgtaat | gattgatttg | 420 |
| 5 | agctaagaat | tttaacagga | ggcaccgccg | gccctagagc | gtaatcacc | ccattccatc | 480 |
| | tttttttaggt | gaaaacatga | attacaataa | taaaatcttg | gtaagtgaat | ctggctcgag | 540 |
| | ccaaaagcac | ctgattcatg | gcgatgaaga | acttttccaa | catgaactga | aaaccatttt | 600 |
| | tgcgcggaac | tggttttttc | tactcatga | tagcctgatt | cctgcccccg | gcgactatgt | 660 |
| | taccgcaaaa | atggggattg | acgaggtcat | cgtctcccg | cagaacgcg | gttcgattcg | 720 |
| 10 | tgcttttctg | aacgtttgcc | ggcatcgtgg | caagacgctg | gtgagcgtgg | aagccggcaa | 780 |
| | tgccaaaggt | tttgtttgca | gctatcacgg | ctggggcttc | ggctccaacg | gtgaactgca | 840 |
| | gagcgttcca | tttgaaaaag | atctgtacgg | cgagtcgctc | aataaaaaat | gtctgggggt | 900 |
| | gaaagaagtc | gctcgcgtgg | agagcttcca | tggttccatc | tacggttgct | tcgaccagga | 960 |
| | ggccccctct | cttatggact | atctgggtga | cgtcgttggt | tacctggaac | ctatgttcaa | 1020 |
| 15 | gcattccggc | ggtttagaac | tggtcggctc | tccaggcaag | gttgtgatca | aggccaactg | 1080 |
| | gaaggcacc | gcggaaaact | ttgtgggaga | tgcataccac | gtgggttgga | cgcacgcgtc | 1140 |
| | ttcgtctcgc | tcgggggag | ctatcttctc | gtcgtcgtc | ggcaatgcgg | cgctaccacc | 1200 |
| | tgaaggcgca | ggcttgcaaa | tgacctccaa | atacggcagc | ggcatgggtg | tgttgtggga | 1260 |
| | cggatattca | ggtgtgcata | gcgcagactt | ggttccggaa | ttgatggcat | tcggaggcgc | 1320 |
| 20 | aaagcaggaa | aggctgaaca | aagaaattgg | cgatgttcgc | gctcggattt | atcgcagcca | 1380 |
| | cctcaactgc | accgttttcc | cgaacaacag | catgctgacc | tgctcgggtg | ttttcaaagt | 1440 |
| | atggaaccgc | atcgacgcaa | acaccaccga | ggtctggacc | tacgccattg | tcgaaaaaga | 1500 |
| | catgcctgag | gatctcaagc | gccgcttggc | cgactctggt | cagcgaacgg | tcgggcctgc | 1560 |
| | tggtctctgg | gaaagcgacg | acaatgacaa | tatggaaaca | gcttcgcaaa | acggcaagaa | 1620 |
| 25 | atatcaatca | agagatagtg | atctgctttc | aaaccttggt | ttcggtgagg | acgtatacgg | 1680 |
| | cgacgcggtc | tatccaggcg | tcgtcggcaa | atcggcgatc | ggcgagacca | gttatcgtgg | 1740 |
| | tttctaccgg | gcttaccagg | cacacgtcag | cagctccaac | tgggctgagt | tcgagcatgc | 1800 |
| | ctctagtact | tggcatactg | aacttacgaa | gactactgat | cgctaacaga | cgagtcgacc | 1860 |
| | atgatgatca | atattcaaga | agacaagctg | gtttccgccc | acgacgccga | agagattctt | 1920 |
| 30 | cgttttctca | attgccacga | ctctgctttg | caacaagaag | ccactacgct | gctgaccagg | 1980 |
| | gaagcgcatt | tggtggacat | tcaggcttac | cgtgcttggt | tagagcactg | cgtgggggtca | 2040 |
| | gaggtgcaat | atcaggtcat | ttcacgcgaa | ctgcgcgcag | cttcagagcg | tcgttataag | 2100 |
| | ctcaatgaag | ccatgaacgt | ttacaacgaa | aattttcagc | aactgaaagt | tcgagttgag | 2160 |
| | catcaactgg | atccgcaaaa | ctggggcaac | agcccggaagc | tgcgctttac | tcgctttatc | 2220 |
| 35 | accaacgtcc | agggcgcaat | ggacgtaaat | gacaaagagc | tacttcacat | ccgctccaac | 2280 |
| | gtcattctgc | accgggcacg | acgtggcaat | caggtcgatg | tcttctacgc | cgccggggaa | 2340 |
| | gataaatgga | aacgtggcga | aggtggagta | cgaaaattgg | tccagcgatt | cgtcgattac | 2400 |
| | ccagagcgca | tacttcagac | gcacaatctg | atgggtctttc | tgtgattcag | tgaccatttt | 2460 |
| | tacaaatggg | cactgcaacc | gcggtcacca | ttaatcaaag | ggaatgtacg | tgtat | 2515 |

<210> 5

<211> 9706.

<212> DNA

<213> Artificial Sequence

5

<220>

<223> A modified DNA sequence

<400> 5

| | | | | | | |
|---------------|-------------|-------------|-------------|-------------|-------------|------|
| 10gctgatttcg | accaattatt | ccacgccgtc | tgccgagca | aggccgtaat | cagttcgtaa | 60 |
| aggggagccc | gcagtaacag | cgaaccgacg | ccggtagctg | cgcgcgtgaa | tcctgagatc | 120 |
| acaaattctt | ggcgattatc | gcatttcaat | ttcacccaag | cgccgttcct | tgcgctgctg | 180 |
| tagcggcggc | catcacgctt | ctccacgatt | ccctccagct | gcacccggca | agcgttggcc | 240 |
| agcagctgca | ctgggtaagc | ttccagggtg | tcgctgaagc | gtacgcgctc | caatgggcag | 300 |
| 15tgctcgagga | gcgcctcaag | tacctggcgg | ctcagttcaa | cgcctctccc | ccgcagatct | 360 |
| gcaccgtcga | ggtagagcag | gtcgaacgcc | acgtacacaa | atttgctcgg | acaccgtgcg | 420 |
| gcgaacgctg | actgcagggc | ttgaaatgct | ggtcgaccgt | cgtcactactg | gaatacgacc | 480 |
| tcaccgtcga | gccaggctga | atggacggct | aggcgtccca | gctattttggc | caggagtggc | 540 |
| atatgatcga | tctagtcaaa | ccggtttttg | gtgaaaagct | gcacctggtc | accatcgatc | 600 |
| 20cgtgctaaca | gacgatacca | gtcgtacttg | atctcgtatc | gccatgctcc | tgccgacgac | 660 |
| agtgtcagta | gagtaaccag | ttgtggcgcg | atccaacacg | ggtttggtgc | cttgggcggg | 720 |
| ttcactatcg | ccacctcgct | ataggaacat | tccttcaggg | tggagtgcgt | aattttctga | 780 |
| aaggggagcc | aggttatgag | tattcacatt | gggtataaac | aacatcactt | atgcgttatt | 840 |
| gacatataac | gtcgtattca | cgattattta | ccatataagt | cttataataa | cgaagccata | 900 |
| 25ttatggaact | cctcatataca | ccgaaaaatc | gcataattcc | cttcagtgcc | ggtgccaacc | 960 |
| ttctggaagt | gcttcgcgag | aacgggtgtag | ctatttccta | cagttgcttg | tctgggcggt | 1020 |
| gcggaacctg | tcgctgccgg | gttatagatg | gcagtgtcat | tgattctggg | gcggaaaatg | 1080 |
| ggcaatcaaa | cctcaccgac | aagcagtatg | tgctcgccctg | tcagtcagta | ctcactggca | 1140 |
| attgcgctat | cgaagtccca | gaagccgacg | aaattgtcac | tcacccggcg | cgaatcatca | 1200 |
| 30agggcacagt | ggtcgcagtc | gagtcgcccc | ctcacgatat | ccgtcgctta | cgcgtacgcc | 1260 |
| tctccaagcc | cttcgagttc | tcacccggac | agtacgcgac | actgcagttc | agccctgagc | 1320 |
| atgcgcgtcc | gtattcaatg | gcaggtttgc | cagatgacca | agaaatggag | ttccacatac | 1380 |
| gcaagggtgcc | gggtgggcgc | gtcacggagt | atgttttcga | acacgtccgc | gaaggtaaaa | 1440 |
| gcacaaagtt | gagcgggcct | cttggtacgg | cttatctacg | tcagaagcac | accggaccga | 1500 |
| 35tgctgtgtgt | agggtggcggg | accggactcg | caccgggtgct | gtcgattgtt | cgcggcgcgc | 1560 |
| tgaagtccgg | tatgacgaac | cccactctcc | tttatttcgg | gggtgcgcagt | cagcaagacc | 1620 |
| tctacgacgc | agagcgattg | cacaaactcg | ccgctgacca | ccctcaactg | accgtacaca | 1680 |
| cggtgattgc | aacgggcccc | attaatgagg | gtcagcgagc | cggcctaatt | accgatgtga | 1740 |
| tcgaaaaaga | catcctttcg | ctggctgggt | ggagggccta | cctgtgcggc | gcaccagcga | 1800 |
| 40tggttggaagc | gttgtgcacc | gtcaccaagc | atcttggaat | atcacccgaa | catattttatg | 1860 |

| | | | | | | |
|--------------|-------------|------------|-------------|------------|-------------|------|
| ccgatgcctt | ctatcccggg | gggatctgaa | tagttcccgg | ccatgcacct | ctgtccatcg | 1920 |
| agaattcatc | aggaagacat | tcaaataaac | gtaaacaata | agggcagcgt | ctgtattttgc | 1980 |
| ggcagcgaaa | tgctccctaa | attcctcatt | taccccatct | gaggattgct | ttatgacagt | 2040 |
| aaagtggatt | gaagcagtcg | ctctttctga | catccttgaa | ggtgacgtcc | tccgctgac | 2100 |
| 5tgctcgagggc | aaggagctgg | cgctgtatga | agttgaaggc | gaaatctacg | ctaccgacaa | 2160 |
| cctgtgcacg | catggttccg | cccgcacgag | tgatgggttat | ctcgagggta | gagaaatcga | 2220 |
| atgccccttg | catcaaggtc | ggtttgacgt | ttgcacaggc | aaagccctgt | gcgcacccgt | 2280 |
| gacacagAAC | atcaaaacat | atccagtcaa | gattgagaac | ctgcgcgtaa | tgattgattt | 2340 |
| gagctaagaa | ttttaacagg | aggcaccctg | ggccctagag | cgtaatcacc | cccatcccat | 2400 |
| 10cttttttagg | tgaaaacatg | aattacaata | ataaaatctt | ggtaagtga | tctggctcga | 2460 |
| gccaaaagca | cctgattcat | ggcgatgaag | aacttttcca | acatgaactg | aaaaccattt | 2520 |
| ttgcgcggaa | ctggcttttt | ctactcatg | atagcctgat | tcctgcccc | ggcgactatg | 2580 |
| ttaccgcaaa | aatgggggatt | gacgaggtca | tcgtctccc | gcagaacgac | ggttcgattc | 2640 |
| gtgcttttct | gaacgtttgc | cggcatcgtg | gcaagacgct | ggtgagcgtg | gaagccggca | 2700 |
| 15atgccaaagg | ttttgtttgc | agctatcacg | gctggggctt | cggctccaac | ggatgaactgc | 2760 |
| agagcgttcc | atttgaaaaa | gatctgtacg | gcgagtcgct | caataaaaaa | tgtctgggg | 2820 |
| tgaaagaagt | cgctcgctg | gagagcttcc | atggcttcat | ctacggttgc | ttcgaccagg | 2880 |
| aggccctcc | tcttatggac | tatctgggtg | acgtcgttg | gtacctggaa | cctatgttca | 2940 |
| agcattccgg | cggttttaga | ctggtcggtc | ctccaggcaa | ggttgtgatc | aaggccaact | 3000 |
| 20ggaaggcacc | cgcggaaaac | tttgtgggag | atgcatacca | cgtgggttgg | acgcacgcgt | 3060 |
| cttcgcttcg | ctcgggggag | tctatcttct | gctcgtctgc | tggcaatgcg | gcgctaccac | 3120 |
| ctgaaggcgc | aggcttgcaa | atgacctcca | aatacggcag | cggcatgggt | gtgttgtggg | 3180 |
| acggatatte | agggtgtgat | agcgcagact | tggttccgga | attgatggca | ttcggaggcg | 3240 |
| caaagcagga | aaggctgaac | aaagaaattg | gcgatgttcg | cgctcggatt | tatcgagcc | 3300 |
| 25acctcaactg | caccgttttc | ccgaacaaca | gcatgctgac | ctgctcgggt | gttttcaaa | 3360 |
| tatggaaccc | gatcgacgca | aacaccaccg | aggctcggac | ctacgccatt | gtcgaaaaag | 3420 |
| acatgcctga | ggatctcaag | cgcgccttgg | ccgactctgt | tcagcgaacg | gtcgggcctg | 3480 |
| ctggcttctg | ggaaagcgac | gacaatgaca | atatggaaac | agcttcgcaa | aacggcaaga | 3540 |
| aatatcaatc | aagagatagt | gatctgcttt | caaacccttg | tttcggtgag | gacgtatacg | 3600 |
| 30gcgacgcggg | ctatccaggc | gtcgtcggca | aatcggcgat | cggcgagacc | agttatcgtg | 3660 |
| gtttctaccg | ggcttaccag | gcacacgtca | gcagctccaa | ctgggctgag | ttcgagcatg | 3720 |
| cctctagtac | ttggcatact | gaacttacga | agactactga | tcgctaacag | acgagtcgac | 3780 |
| catgatgatc | aatattcaag | aagacaagct | ggtttccgcc | cacgacgccg | aagagattct | 3840 |
| tcgtttcttc | aattgccacg | actctgcttt | gcaacaagaa | gccactacgc | tgctgaccca | 3900 |
| 35ggaagcgcac | ttgttggaac | ttcaggctta | ccgtgcttgg | ttagagcact | gcgtggggctc | 3960 |
| agaggtgcaa | tatcagggtca | tttcacgcga | actgcgcgca | gcttcagagc | gtcggtataa | 4020 |
| gctcaatgaa | gccatgaacg | tttacaacga | aaattttcag | caactgaaag | ttcgagttga | 4080 |
| gcatcaactg | gatccgcaaa | actggggcaa | cagcccgaag | ctgcgcttta | ctcgctttat | 4140 |
| caccaacgtc | caggccgcaa | tgagcgtaaa | tgacaaagag | ctacttcaca | tccgctccaa | 4200 |
| 40cgtcattctg | caccgggcac | gacgtggcaa | tcaggctgat | gtcttctacg | ccgcccggga | 4260 |

agataaatgg aaacgtggcg aaggtggagt acgaaaattg gtccagcgat tcgtcgatta 4320
 cccagagcgc atacttcaga cgcacaatct gatggtcttt ctgtgattca gtgaccattt 4380
 ttacaaatgg tcaactgcaac cgcggtcacc attaatcaaa gggaatgtac gtgtatgggc 4440
 aatcaacaag tcgtttcgat aaccgggtgca ggctcaggaa tcggtctcga actggttcgg 4500
 5tcctttaagt cggccgggta ttacgtatcc gctctcgtac gaaacgagga gcaagaggcg 4560
 cttctttgca aagagttaa ggacgcactc gagattgtag tgggcgatgt ccgggaccac 4620
 gcaacaaatg agaagctgat aaagcaaca atcgatagat tcggtcatct tgattgtttt 4680
 attgcaaatg ccggtatctg ggattacatg ctgagcatcg aagagccttg ggagaaaata 4740
 tcgagcagtt ttgacgaaat attcgacatt aatgtcaaga gctatttcag tggcatcagt 4800
 10gccgccctgc cggaactgaa aaagactaac ggatcagtggt tgatgaccgc ttcggtgtcg 4860
 tcccatgcgg tcggtggtgg tggttcttgc tacatcgcca gcaagcatgc ggtgctcggg 4920
 atggttaagg ctttggccta cgaattggcc cccgaagttc gcgtgaacgc tgtttcgccg 4980
 gggggcaccg tgacgtctct gtgcgggttc gcgagcgccg gtttcgacaa aatgcacatg 5040
 aaagacatgc ccggcatcga cgatatgatc aaagggtctc cgctcttggt gtttgacgcc 5100
 15aagcccgaag acgtggtggc accctatttg ttgctggctt cgcgaaagca aggaaaattc 5160
 atcaccggca ccgtgattag cattgatggc ggtatggcgc tcggtcgcaa gtgagcttgt 5220
 agccgatcag aagtataga cacatttcag gtgacgcccc atgaagacaa aactgtttat 5280
 caataacgcc tggatcgatt ctagtgacca gcagaccttc gagcgcatat acccgcgcag 5340
 cagcgatgtg gtgactgaga gcgcaaacgc cacagtgcg gacgcgataa agggggcgca 5400
 20agcgcccgag gaggcgttca agacctgga ggccgttga ccttcagagc gtcgccgcct 5460
 tctcctaaag gtcccgatg tcatggaaag taaaacaccc aagttcatcg aagtgatggc 5520
 catggagggt ggagcttcg ccctttgggc cggattcaac gtccatgctg ctgccaatgt 5580
 gttccgagag gctgcctcgc tggctacca aattcagggt gaaaccatcc caacggacaa 5640
 agccgaaaac ctctcaatga cactacgtca gccggtcggc ccgataccta gcatcgctcc 5700
 25atggaacggc accgcagtgc ttgcggcacg agccatcgct tatccgctgg tctgtggcaa 5760
 cactgtggtg ttcaaaagct ctgaatttag tcccgcgacg catgccctga tcaccagtg 5820
 cgtgcaggaa gccgggctgc ccgctggcgt gctcaattac ctcaactctt cgctgaccg 5880
 ttcgcccag atcgtgacg cactgatctc tgccaaggag atccgccga tcaactcac 5940
 gggttccacc cgctggggca gcattatcgc gcagaaagcc gcgcaacacc tcaagcgtg 6000
 30cctgctggag ctccggcgca agtccccgct tattgttctg gatgatgcag acatcgatgc 6060
 ggcggtcaag gcagcgggtg tcggtagctt cctgttccaa ggtcagatct gcatgtccac 6120
 tgagcgcttg atcgttgatg agaagatagc cgacgaattt gtcgcaaaat ttgtcgaaaa 6180
 aactaagcgc ttgagcgag gcgacccgtg cgtaactggc gactgcatca tcggcccgat 6240
 ggtctcgcca aattcgggtg agcggatcaa tggtttgttc aaagacgca tcgacaaagg 6300
 35ggcaaaagtt gtttcggcg gcttggccca aggtgcgctc atgccggcca cgatcctgga 6360
 tcacgtcaaa tctgacatgc ggatttacga tgaggagacc tttgggtcca tcaccgtggt 6420
 aatccgttgt aaaggcgaag cagaggcgt ccgcattgcc aacgacagcg tctatggcct 6480
 gtcgtcgggc gtatttggcc gcgacatcaa ccgcgtctca cgcgtgggta tgtccatcga 6540
 atatggttct gtacacatca acggttcgac cgtccagaac gaggcgcagg ctccctacgg 6600
 40aggcaccaag aacaccggct acgggcgctt cgacggcgt gctgtaatcg acgagttcac 6660

agagatcaag tggctgacca tcgaaccttt cgagcagcaa tatcccttct gataagcact 6720
 aactcccagg aatcaaacta tgagtaagca agctgcagtt atcgagctcg gatacatggg 6780
 tatctcggtc aaggacctg atgcgtggaa atcatttgcc acggatatgc taggtctgca 6840
 agttcttgat gaggtgaga aggaccgttt ctatctgcgg atggattact ggcacatcg 6900
 5gatcgtagtc catcacaacg gacaggacga cttggagtac ctaggctggc gtgtagccgg 6960
 caagccggag ttcgaagctc tgggtcaaaa gcttattgat gccggttaca agatccgcat 7020
 ctgcgacaaa gttgaggctc aggagcgtat ggtgttgggt ctgatgaaga cagaagatcc 7080
 gggcggcaac ccgaccgaga tattctgggg ccccggtatc gacatgagca acccggtcca 7140
 tcccggtcgc cccctgcacg gaaagtttgt gaccgggtgac caaggcttgg gccattgcat 7200
 10cgttcgccaa accgacgtcg cagaagctca taagttttat agcctgctgg gcttccgtgg 7260
 ggacgtcgaa taccggattc cgttgcccaa cggcatgact gccgaactgt cgttcatgca 7320
 ttgcaacgcc cgtgatcact ccattgcttt tgggtgccatg cccgctgcca aacgactcaa 7380
 tcacttgatg cttgagtaca cccatatgga agacttggga tacacgcacc aacagtttgt 7440
 aaagaacgaa attgacattg ccttgacgtc tggcattcac gccaacgaca aggcgttgac 7500
 15gttctatggt gcaacgcctt cgggctggct cattgagccc ggctggcgag gtgccacggc 7560
 catagatgaa gcggagtatt acgtcggcga catcttcggc catggcgtgg aggccactgg 7620
 atatggcctg gatgtaaaac tgagctaaag atgcgcgtc gttggggcag gctctagtc 7680
 agcatcttca tacgcaacca accttgacgg gcgatgagat caaaggacgt taaagcgaag 7740
 ggggaagtgg tccggccatg cgcataccga tccatgacat ttgtttcata gtatataggt 7800
 20agataggtga atcaagcgt tagtcaacta gtggacacat ctgttccatg aggtatctta 7860
 ctatctattc aaaacaagaa taataaatag gatgaaaata ataatagataa aaagaacgat 7920
 ttgtcttggt tatctcttat tctgtttggc aagccccaca tgggcccgaag agtcgccttg 7980
 gacgtaccgt attggtatga ctaatgtagc tttcgatgct agcgcaaaag tatacttaaa 8040
 tggtcagcgg gtgccaggag gaagcgtga tgcgagcgat aacaacgcgc ttacattcga 8100
 25cttcggctac gccatcaacg accagtgga tgtacgtgcg attgtcggta ttccgcctac 8160
 aactaaagt acgggcgcag gcacacttcc tggtatccag ctggggaaaa taacttacgc 8220
 tccaacagta ttaacgttga actataacct ccccgctttg ggtcccgttc gccctcacat 8280
 aggtgcggga gtcaattaca cgcggatttt tgaaagtcgg gacgctaata taaaatcgtt 8340
 cgatgccgac cacgcttggg ccccgcgct acatgttggg gccgatattg acgttaaccg 8400
 30tggttggttc gttagcattg atatccggaa gttatacctg aaaagcgacg catcagggtta 8460
 cttggggcca caggaggcta aagcacgggt aactcttgac ccattactaa cttcgatcgc 8520
 gatcggacgc caattctgat gattctgttt aaagtcttt atctatctaa ccgcaaaggg 8580
 tgtttccatg tcgaataaaa ttatgaaaac gtcgcgtctt accgccgaag atatcaacgg 8640
 cgcctggact ataatgccc caccctcgac gcctgatgct tctgattggc gcagcactgc 8700
 35caccgtggac ttagaagaga ctgcccgcac agttgaagag ctgattgcag ctggtgtcaa 8760
 cggatttcta agtatgggtta cttttggtga gtgcgccacg ttgacctggg atgaaaaacg 8820
 tgattatgtc tcgacgattg tcgagaccat tcgtggtcgc gtgccttatt tctgtggcac 8880
 gacagcctta aataccggag aagtcacccg ccagaccga gagcttatcg atattggcgc 8940
 caacggcact atgctcgggg tgccgatgtg ggtgaagatg gacctgccta cagcggttca 9000
 40gttctatcgt gatgttgacg atgcggtacc agaggctgcc attgcgattt acgccaaccc 9060

```

cgaagcattc aagttcgact tccctcgccc attctgggca gagatgtcca aaattccgca 9120
ggtagtgact gcgaagtatc taggcacgga aatgcttgac ttggacctga gactggcacc 9180
caacatccgc ttccttcccc acgaagatga ctattacgag gccgcacgca tcaatccgga 9240
gcgcataacc gcgttctggt caagcggggc catgtgcggc ccggctaccg ccacatggtt 9300
5gcgtgacgaa gtggtgcggg ccaagagcac cggtgactgg gccaaaggcca aagccatctc 9360
cgatgatatg cgtgcagccg actcgacatt gtttcgcgct ggcgactttt cggagttctc 9420
gaagtataat atcgggcttg aaaaggcacg gatggatgag gctgggtggc tcaaggctgg 9480
gccctgccgt ccgccctaca acctgtgtcc agaagactac ctgcctggtg cacagaaatc 9540
aggcaaggct tgggcgcgag tgcacgctaa atacagtaat gaattgaagt agttcacctc 9600
10cgcgagacct agtgacaggg tggcgagagc gctgaggggt caggaattaa gtgagctaaa 9660
gcacatttct tgcgccaggc attgccagat cagcaaagtt tgctga 9706

```

```

15      <210> 6
          <211> 2294
          <212> DNA
          <213> Artificial Sequence

20      <220>
          <223> A modified DNA sequence

```

```

          <221> unsure
          <222> (186) ... (186)

25      <223> n = a or t or g or c

```

```

          <400> 6

agggcagcct ctgtatttgc ggcagcgaaa tgctccctaa attcctcatt taccatctct 60
gaggattgct ttatgacagt aaagtggatt gaagcagtcg ctctttctga catccttgaa 120
30 ggtgacgtcc tcggcgtgac tgctgagggc aaggagctgg cgctgtatga agttgaaggc 180
gaaatntacg ctaccgacaa cctgtgcacg catggttccg cccgcatgag tgatgggttat 240
ctcgagggta gagaaatcga atgccccttg catcaaggtc ggtttgacgt ttgcacaggc 300
aaagccctgt gcgcacccgt gacacagaa atcaaaacat atccagtcaa gattgagaa 360
ctgcgcgtaa tgattgattt gagctaagaa ttttaacagg aggcaccccg ggccctagag 420
35 cgtaatcacc cccattccat ctttttttagg tgaaaacatg aattacaata ataaaatctt 480
ggtaagtga tttggtctga gccaaaagca cctgattcat ggcgatgaag aacttttcca 540
acatgaactg aaaaccattt ttgcgcggaa ctggcttttt ctactcatg atagcctgat 600
tcctgccccg ggcgactatg ttaccgcaaa aatggggatt gacgaggtca tcgtctcccg 660
gcagaacgac gggttcgattc gtgcttttct gaacggttgc cggcatcgtg gcaagacgct 720
40 ggtgagcgtg gaagccggca atgccaaagg ttttggttgc agctatcacg gctggggctt 780

```

cggctccaac ggtgaactgc agagcggtcc atttgaaaaa gatctgtacg gcgagtcgct 840
 caataaaaaa tgtctggggt tgaagaagt cgctcgctg gagagcttcc atggcttcat 900
 ctacgggtgc ttcgaccagg agggccctcc tcttatggac tatctgggtg acgctgcttg 960
 gtacctggaa cctatgttca agcattccgg cggtttagaa ctggtcggtc ctccaggcaa 1020
 5 ggttgtgatac aaggccaact ggaaggcacc cgcgaaaaac tttgtgggag atgcatacca 1080
 cgtgggttg acgcacgcgt ctctgcttcg ctcgggggag tctatcttct cgtcgctcgc 1140
 tggcaatgag gcgctaccac ctgaaggcgc aggccttgcaa atgacctcca aatacggcag 1200
 cggcatgggt gtgttgtggg acggatattc aggtgtgcat agcgcagact tggttccgga 1260
 attgatggca ttcggaggcg caaagcagga aaggctgaac aaagaaattg gcgatgttcg 1320
 10 cgctcgatt tatcgagcc acctcaactg caccgttttc ccgaacaaca gcgatctgac 1380
 ctgctccggt gttttcaaag tatggaacct gatcgacgca aacaccaccg aggtctggac 1440
 ctacgccatt gtcgaaaaag acatgcctga ggatctcaag cgccgcttgg ccgactctgt 1500
 tcagcgaacg gtcgggcctg ctggcttctg ggaaagcgac gacaatgaca atatggaaac 1560
 agcttcgcaa aacggcaaga aatatcaatc aagagatagt gatctgcttt caaaccttgg 1620
 15 tttcgggtgag gacgtatacg gcgacgcggt ctatccaggc gtcgtcggca aatcgcgat 1680
 cggcgagacc agttatcgtg gtttctaccg ggcttaccga gcacacgtca gcagctccaa 1740
 ctgggctgag ttcgagcatg cctctagtag ttggcatact gaacttacga agactactga 1800
 tcgctaacag acgagtcgac catgatgata aatattcaag aagacaagct ggtttccgcc 1860
 cacgacgccg aagagattct tcgtttcttc aattgccacg actctgcttt gcaacaagaa 1920
 20 gccactacgc tgcagacca ggaagcgcat ttgttgaca ttcaggctta ccgtgcttgg 1980
 ttagagcact gcgtggggtc agaggtgcaa tatcaggctca tttcacgca actgcgcgca 2040
 gcttcagagc gtcgttataa gctcaatgaa gccatgaacg tttacaacga aaattttcag 2100
 caactgaaag ttcgagttga gcatcaactg gatccgcaaa actggggcaa cagcccgaag 2160
 ctgcgcttta ctgcctttat caccaacgct caggccgcaa tggacgtaaa tgacaaagag 2220
 25 ctacttcaca tccgctccaa cgctattctg caccgggcac gacgtggcaa tcaggtcgat 2280
 gtcttctacg ccgc 2294

<210> 7

<211> 4355

30 <212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

35

<400> 7

atcgataaaa ccatgaggggt aataaataat gatgaaggag actatctaga attcgttctg 60
 gagacggtgc cggtagcata accgtagacc agcaggctga gcaaaatacc tcggggtgat 120
 aagccgcgct accgcgtccg ccatatgcct gagtgagtg gctaaggctg ggccgctcaa 180
 40 tggcttcgag aataaagcgg gccagggtgat cttctgtcaa tctcgccgaa tgggtggaag 240

| | | | | | | | |
|----|-------------|------------|-------------|------------|-------------|-------------|------|
| | cacatcaacc | tgaccggcga | ttacgtctgg | cggcagagcc | gcagactgga | ggtcgggaag | 300 |
| | ttccggccgt | tacggtggcc | ctgaaaacct | tagtgtagca | ttttttccgt | cttctctatt | 360 |
| | cggcctgcgt | tgcgacgttc | gatctgacta | gggcgcagcg | gcaacatcgg | ctgagtagca | 420 |
| | tccagcgcct | ggatctgggt | tatttcgtca | accgacaaca | ccagagcggt | atccgggtggg | 480 |
| 5 | ttcaggtaaa | gccaaccac | attgactacc | tctcggcaaa | atgagggtcg | ttactaattt | 540 |
| | tgaacgtctt | caagcgatgg | ggtttttaggt | cgaagctgc | ccagacctgc | ccaacctacc | 600 |
| | acgtcgtgac | ccgagcatat | tcggcccgcg | gcggcaaaaa | acacagaaat | gagcgggtga | 660 |
| | cccgatcgcc | tttgatcgat | tctccgcttt | caaaatgggc | gggggctgaa | gtcagccaga | 720 |
| | aatacatgac | tacttcagac | gattcgcagt | attcacgctg | gtgataaaca | aattcaacta | 780 |
| 10 | tgctttattg | acaaataaaa | gcacgctcac | catcatcgcg | aatacaaatc | ttataaaaaat | 840 |
| | taagccggat | ttatggaact | tctcatacag | ccaaacaatc | gcctcattag | ctttagtcct | 900 |
| | ggcgccaacc | ttctggaagt | gcttcgcgaa | aacggtgtcg | ctatttccta | cagttgtatg | 960 |
| | tctgggctgt | gcggaacctg | ccgctgtcgg | gttacagatg | gcagtgtaat | tgattcgggg | 1020 |
| | gcgggaagcg | ggttaccaa | cctcgtggac | gagcattatg | tgctcgctcg | tcagtcagta | 1080 |
| 15 | cttactcaca | attgcgcgat | cgaaatccca | gaaaccgacg | aaatcgtcac | ccaccggcg | 1140 |
| | agaatcatca | agggcactgt | ggtcgccgtc | gagtcgcca | ctcacgatat | ccgtcgcccta | 1200 |
| | cgcgtacgcc | tcgctaagcc | cttcgagttc | tcaccgggac | agtagcgac | attgcagttc | 1260 |
| | agtcctgagc | atgcgcgtcc | gtattcaatg | gcaggctctg | cagatgacca | agaaatggag | 1320 |
| | ttccacatac | gcaagggtgc | gggtgggcgc | gtaacggagt | atgttttcga | gcacgtccgc | 1380 |
| 20 | gaagggtacaa | gcatcaagtt | gagcggggcca | cttggtacgg | cttatttgcg | tcagaaccac | 1440 |
| | accggggccga | tgctctgtgt | gggcgggtggg | accggactag | caccgggtgct | gtcgattggt | 1500 |
| | cgcgggcgcg | tgaagttggg | tatgacaaac | cccatcctcc | tttatttcgg | agtgcgcagt | 1560 |
| | cagcaagacc | tctacgacgc | agagcgattg | cacaaactcg | ccgctgatca | ccctcaactg | 1620 |
| | accgtacaca | cggtaatcgc | aatggggccc | attaatgaga | gtcagcgagc | cggctctagtt | 1680 |
| 25 | accgatgtga | tcgaaaaaga | catcatttcg | ctgggtgggt | ggagggccta | cctgtgcggc | 1740 |
| | gcaccagcga | tggttgaaag | gctttgcacc | gttaccaagc | atcttggaat | atcaccgaa | 1800 |
| | catatttatg | ccgatgcctt | ctatcccggg | ggaatctgaa | tagtcccttt | ccctgcacct | 1860 |
| | ctgtccatcg | aggactcatc | aggaggatac | tcaaataagc | gcaaataata | acagtcgcgt | 1920 |
| | cagtatttgc | ggcagcgaaa | tggtttccct | ctccctcatt | taccccatct | gaggatagtt | 1980 |
| 30 | ttatgacaga | aaaatggatt | gaagcagtcg | ccctttctga | catcccagaa | ggtgatgtcc | 2040 |
| | tcggcgtgac | tgtagagggc | aaggagctgg | cgttgtacga | agtggaaagg | gaaatctacg | 2100 |
| | ctaccgacaa | cctgtgcacg | catggtgcgg | ccgcatgag | cgatgggttat | ctcgaggggc | 2160 |
| | gagaaatcga | atgcccttg | catcaaggtc | ggtttgacgt | ttgtacaggc | agagccctct | 2220 |
| | gcgcccccg | gacagagaac | atcaaaacat | atgcagtcaa | gattgagaac | ctgcgcgtaa | 2280 |
| 35 | tgattgattt | gagcggagag | ttttaatagg | aggtacccc | gaccctagcg | cgtaactacc | 2340 |
| | ccaattccat | cttttttagg | tgaaaacatg | aattacaaaa | acaaaatatt | ggtgagtga | 2400 |
| | tctgggctga | cccaaaagca | cctgattcat | ggcgtgaag | aacttttcca | gcacgaactg | 2460 |
| | agaaccattt | ttgcgcggaa | ctggcttttt | ctactcatg | acagcctgat | tccatcccc | 2520 |
| | ggcgactatg | ttaccgcaaa | aatgggtatt | gacgaggtca | tcgtctctcg | gcagagcgac | 2580 |
| 40 | ggttcgattc | gtgccttcct | gaacgtttgt | cggcaccgtg | gcaagacgct | ggttaacgg | 2640 |

gaagccggca atgccaaagg tttcgtttgc agctatcacg gctggggctt cggctccaac 2700
 ggtgaactgc agagcgttcc attcgaaaaa gagctgtacg gcgagtcgct caacaaaaaa 2760
 tgtctggggt tgaagaagt cgctcgctg gagagcttcc atggcttcat ctatggttgc 2820
 ttcgatcagg aggccctcc tcttatggac tatctgggtg acgctgcttg gtacctagag 2880
 5 cccatcttca aacattcagg cggtttagaa ctggtcggtc ctccaggcaa ggttgtgatc 2940
 aaggccaact ggaaggcacc cgcggaaaac tttgtggggg atgcatacca cgtgggttgg 3000
 acgcacgcgt cttcgcttcg ctccggagag tctatcttcg cgctcgctcg tggcaacgca 3060
 gtgctgcccc ctgaagggtg aggcttgcaa atgacctcca aatacggcag cggcatgggt 3120
 gtgttgtggg acggatatc aggcgtgcat agcgcagact tggttccgga gttgatggca 3180
 10 ttcggcggct ctaagcagga aaggctgaac aaagaaattg gcgatgttcg cgcccgatt 3240
 tatcgacgcc acctcaactg caccgttttc ccgaacaaca gcatgctgac ctgctcgggt 3300
 gttttcaaaag tatggaaccc gatcgacgca aacaccaccg aggtctggac ctacgccatt 3360
 gtcgaaaaag acatgcccga ggatctcaag cgccgcttgg ccgactcggg tcagcgtacg 3420
 gttgggcctg ctggcttctg ggaaagcgac gacaatgaca atatggaaac agcgtcgcaa 3480
 15 aacggcaaga aatatcaatc cagagatagt gatctgcttt caaaccttgg tttcggttaag 3540
 gacgtatacg gcgacgcggg ctatcctggc gtcgtcggca aatcggcgat cggcgagacc 3600
 agttatcgtg gtttctaccg ggcttaccag gcacacgtca gcagctccaa ctgggctgag 3660
 ttcgaggatg cctctagtac ttggcatacc gaactgacga agactactga tcgctaacag 3720
 acgagtcgac catgatgatc aatattcaag aagacaagct ggtctccgcc cagcagccc 3780
 20 aagagtttct tcgtttcttc aattgccacg acgcggcttt gcaacaagaa gccaccacgc 3840
 tgctgaaccg ggaagcgcat ctgttggaac ttcaggctta ccgggcttgg ttagagcact 3900
 gcgtggggtc agaggttcaa tatcagggtc tttcacgca actgcgcgcc gcttccgagc 3960
 gccgttataa gctcaatgaa gccatgaacg tttacaacga aaattttcag caactgaaag 4020
 ttcgaatcga gcatcaactg gatccgcaaa actggagcaa cagcccgaag ctgcgcttta 4080
 25 ctgcgttcat caccaatgtc caggccgcaa gggacgtaga tgacgaagag ctacttcaca 4140
 tccgtcccaa cgtcattctg caccgggcac gacgtggcaa tcaggctgat gtcttctacg 4200
 ccgcccggga agacaaatgg aaacgtggcg aagggtggagt gcgaaaattg gtccagcgat 4260
 tcgtggatta ccagagcgc atacttcaga cgcacaatct gatggtcttt ctgtgatcca 4320
 gtgaccactt ttacaaatgg tgactgctac cgcgg 4355

30

<210> 8

<211> 2176

<212> DNA

<213> Artificial Sequence

35

<220>

<223> A modified DNA sequence

<400> 8

40 gtgatgtcct cggcgtgact gtcgagggtg aggatctggc actgtacgaa gtggaaggcg

60

| | | | | | | | |
|----|-------------|------------|-------------|------------|------------|------------|------|
| | aaatctacgc | taccgacgac | ctgtgcacgc | atggtgccgc | ccgcatgagc | gatgggtatc | 120 |
| | tcgaggggcg | agaaatcgaa | tgccccctgc | atcaaggctg | gtttgacggt | tgtacaggca | 180 |
| | gagccctctg | cgccccctg | acagagaaca | tcaaaacata | tgcagtcaag | attgagaacc | 240 |
| | tgcgcgtaat | gattgattta | agctgagaat | ttttaatagg | aggcgccccg | gaccatagag | 300 |
| 5 | cgtaattatc | cccattccat | cttttttttag | gtgaaaacat | gaattacaaa | aacaaaatct | 360 |
| | tggttaagtga | gtctgggctg | acccaaaagc | acctgattca | tggcgatgaa | gaacttttcc | 420 |
| | agcacgaact | gagaaccatt | tktgcgcgga | actggctttt | tctcactcat | gacagcctga | 480 |
| | ttccttcccc | cggcgactat | gttaccgcaa | aaatggggat | tgacgaggtc | atcgctcttc | 540 |
| | ggcaaagcga | cggttcgatt | cgtgccttcc | tgaacgtttg | ccggcaccgc | ggcaagacac | 600 |
| 10 | tggtgaacgc | ggaagccggc | aatgctaaag | gtttcgtttg | cagttatcac | ggctggggct | 660 |
| | tcggtctcaa | cgcggaactg | cagagcgctt | cattcgaaaa | agagctgtac | ggcgagtcgc | 720 |
| | tcaacaaaaa | atgtctgggg | ttgaaagaag | tcgctcgctg | agagagcttc | catgggttca | 780 |
| | tctatggttg | cttcgatcag | gaggccccct | ctcttatgga | ctatctcggt | gacgctgctt | 840 |
| | ggtacctgga | acccatcttc | aaacattcag | gcggtttaga | actggtaggt | cctccaggta | 900 |
| 15 | aggttgtgat | caaggccaac | tggaaggcac | ctgcggaaaa | ctttgtgggt | gatgcatacc | 960 |
| | acgtcggttg | gacgcacgcg | tcttcgcttc | gcacaggcga | atctatcttc | tcgtcgctcg | 1020 |
| | ctggcaacgc | agtgetgccg | cctgaagggt | cgggcttgca | aatgacctcc | aaatacggca | 1080 |
| | gcggtatggg | tgtgttggtg | gacggatatt | caggcgctga | tagcgcagac | ttggttcggg | 1140 |
| | aattgatggc | attcggcgcc | gctaagcagg | aaaggctgaa | caaagaaatt | ggcgatgttc | 1200 |
| 20 | ccgcccggat | ttatcgcagc | cacctcaact | gcaccgtttt | cccgaacaac | agcgtgctga | 1260 |
| | cctgctcggg | tgtttttcaa | gtatggaacc | cgatcgacgc | aaacaccacc | gaggctcggg | 1320 |
| | cctacgccat | tgctgaaaaa | gacatgcccg | aggatctcaa | gcgccgcttg | gccgacgcgg | 1380 |
| | ttcagcgaac | ggtcgggcct | gctggcttct | gggaaagcga | cgacaatgac | aatatggaaa | 1440 |
| | cagcatcgca | aaacggcaaa | aaatatcaat | ccagagatag | tgatctgatt | tccaaccttg | 1500 |
| 25 | gtttcgggaa | ggatgtatac | ggcgacgcgg | tctatcctgg | cgctcgtcga | aaatcggcga | 1560 |
| | tcggcgagac | cagttatcgt | ggtttctacc | gggcttacca | ggcacacgtc | agcagctcca | 1620 |
| | actgggctga | gttcgaggat | gcctctagta | cttggcatac | cgaactgacg | aagactactg | 1680 |
| | atcgctaaca | gacgagtcga | ccatgatgat | caatactcaa | gaagacaagc | tggtatccgc | 1740 |
| | ccatgacgcc | gaagagtttc | ttcgtttctt | caattgccac | gactcggctt | tgcaacaaga | 1800 |
| 30 | agccaccacg | ctgctgacct | gggaagcgca | tctgctggac | attcaggctt | accggacttg | 1860 |
| | gtagagcac | tgctgggggt | cagaggttca | atatcaagtc | atttcacgcg | aactgcgcgc | 1920 |
| | cgttcccgag | cgccgttata | agctcaatga | agccatgaac | gtttacaacg | aaaatttcca | 1980 |
| | gcaactgaaa | gttcgagtcg | agcatcaact | ggattcacaa | aactggagca | acagcccga | 2040 |
| | gctgcgcttt | actcgcttca | tcaccaatgt | ccaggccgca | atggacgtaa | atgatgaaga | 2100 |
| 35 | tctgcttcac | gtccgctcca | acgtcgttct | gcaccgggca | cgacgtggca | atcaagtcga | 2160 |
| | tgtcttctac | gccgcc | | | | | 2176 |

<210> 9

<211> 14462

40

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

5

<400> 9

| | | | | | | |
|--------------|-------------|------------|-------------|------------|-------------|------|
| gtcgactccg | gtggccaaat | cgacgggagc | acaagcggtta | cagcatcggg | gctatcggcc | 60 |
| tcaaccgctg | agacgcatct | acatacccaa | aagcaatggc | aagaagcgtc | cgctgggaat | 120 |
| tccaacgatg | cgggatcgtg | cgatgcaggc | gctgtggaaa | ctggcacttg | agccggttgc | 180 |
| 10cgaaaacgct | gtagatccca | actcctatgg | atttcggccg | cagcgatcca | ctgccgatgc | 240 |
| aatcgacat | tgcttcaatg | cactggcgaa | acgtggttcg | gcgcactggg | tactggaagg | 300 |
| tgacattcga | ggctgtttcg | acaatatcag | tcacgattgg | ctgctcacca | acgtgcccac | 360 |
| ggataagggtg | gttctgcgca | aatggcttcg | agcgggggat | gtcgatcagg | gagccttggt | 420 |
| cgcaacggag | gcaggaaccc | cgcaaggggg | aatcatctct | ccggtacttg | cgaattggac | 480 |
| 15gctggatggc | ctggaagatg | tcgtccatgc | aagcgtggct | tcgacagcgc | gcaagcgtaa | 540 |
| gccattcaag | atacacgtcg | tacgatatgc | cgatgacttc | atcatcacag | gggccacgca | 600 |
| agctgttctg | caacatcagg | ttcgtcctgc | aattgaggcg | ttcctgaaac | agcgtggact | 660 |
| ggaactctct | gatgaagaga | ctcagataac | gcatactctca | caaggcttcg | attttctggg | 720 |
| ccagaacgta | cgcaagtacg | ccggcaagct | actcatcact | ccggcttgca | agagtgtgaa | 780 |
| 20ggcactgctg | gataaggtgc | gggaaatcgc | gaacgcaaac | aaggcggcga | ctcaagctaa | 840 |
| cttgatcttg | accctgaacc | cggtaatccg | aggatggggc | atgtatcatc | gccatgttgt | 900 |
| cgccgccaac | gtttcgcgatg | gatcgaccat | cagatttggc | aagtgttatg | gcgctgggca | 960 |
| gttcgtcggc | acgccatgaa | aaatgcctat | tgggtaaaac | aacgctactt | tcgtgtcgta | 1020 |
| ggtcaacggc | actgggtttt | cgccactcag | gagaaagcgc | gcggcttgag | tcaaccgcct | 1080 |
| 25tggtgtgatg | ccgccgccag | tgtttcgatc | gtgcggcata | tcaaaatatg | cagtgcagcg | 1140 |
| aaccggttcg | atccggcatg | gacgttttac | cttgagcgtc | gaagagcgca | tcgtcaggtt | 1200 |
| gcccagtctt | actcaggctg | ctggaaggct | tgagccgat | gaggggtgac | tcccaagtac | 1260 |
| ggttcttagg | ggagacaggg | gtggtaacac | ccctgtttta | cccgaactga | agcgcatatc | 1320 |
| ggagccgacg | cagaatcggg | cctggtacat | catgtccatg | cacagccgcc | aacgtggctg | 1380 |
| 30acgtgacgga | agttgcccac | ctgttgacag | gtggcgagaa | cgttgtctgt | gcttgggggt | 1440 |
| acaccggcgt | agagaaacgg | cctgagcacg | acggtcggca | agtgatctgg | cagatctgcg | 1500 |
| cgacgcagca | cctacaagca | tttgagcaaa | cgagtgccg | cgtacaaggc | caggcgcaag | 1560 |
| atcgagaagg | caaagccca | ggtacgtgcg | aaggtcgaac | atccgtttcg | ggtgatcaag | 1620 |
| cgtcagttcg | gttatgtgaa | gaccgccttc | cgtgtcctgg | ccaagaatac | ggcggaattg | 1680 |
| 35accacactgt | tcgcgtgtc | gaactctgtg | gatgggtgcg | cggcaattat | tgctgtctgc | 1740 |
| gggagagggtg | cgcccgtgag | tgacagaaaa | ccagggtttt | gcctcggtta | tccaataaca | 1800 |
| cagcgactga | aaatcgggca | tttcggcatc | cctaagccgc | ccatttccga | ctgatgagca | 1860 |
| acttgttcgg | agtttcccta | aagcacaagc | gcaagccatc | agaaaagagt | gcctgagtggt | 1920 |
| gctccaagat | agcccgaac | taaggccgga | catggaaaag | accaagatca | cccatgtgaa | 1980 |
| 40tgatggactt | gtttttgtgg | gccatcggat | cattcgcaag | cgtagccgat | acggagatat | 2040 |

| | | | | | | |
|---------------|------------|-------------|-------------|-------------|-------------|------|
| gcgtgtgatc | acgacgatcc | gagacaaagc | cagaagattc | gcagcctcac | tgacggcgct | 2100 |
| gctatcaggt | aaccacagtg | aaagtaagat | cgacatgggt | gaatcagtc | accgaaagct | 2160 |
| taaaggctgg | gcggcggtta | tcaattcggt | gattcaaagc | caaagtattc | agctatatcg | 2220 |
| atcgtgtcgt | gttttgga | ctggccatt | ggctggctca | aaaataccgc | tcccagggtt | 2280 |
| 5gtcttgctg | ttcctggccc | gccgatgaac | accacattct | gcgcgggtatc | tgtgaactcc | 2340 |
| aggttggacc | gctccctatg | ggcatcaaca | atatctcggt | gccttggctg | ggatacgcaa | 2400 |
| ctaaccaaaa | tgcgcgactg | gcgcagtcag | cttccatag | ggagacgctg | gaagcgctcg | 2460 |
| aatagaccgt | caatcagaag | accttctctg | ctcccgctga | actgaaagcc | tcaccgagtg | 2520 |
| cgaacgtcaa | cgggaaaggc | gcgcaaagac | ttgccttgc | cacctcgctca | cgagcggtaca | 2580 |
| 10aacgctgtga | gctgggtgta | cgggtgcgct | gcttagcggt | ggtttttccg | ttctgtgagc | 2640 |
| caaccccgctg | atatcgtagt | gcaggaccat | gccgttaagg | agggctgcct | ggtcagcatt | 2700 |
| ttcagccagc | gccggaagtt | cctgcacttc | ggccttgaac | tggcggtcca | atgcctggat | 2760 |
| atggccatag | ggaggtacgc | tttggcgtaa | ggcggtgcat | gcagcttggg | gctgtcgagg | 2820 |
| ctgatgggtg | ggagcttgag | cagcttcatt | tgcctgcca | actccaccgc | ctgaagggaag | 2880 |
| 15ttcccgccca | agtcacgag | gaagcgctgc | cggaaagctgg | ccgggggtgac | tcttgcggcg | 2940 |
| agggtggcgaa | aagcgagcga | gtcgtaactg | gcgcgctcaa | tcctggagct | gggaactttg | 3000 |
| ccggtggcat | aaccatagac | catcgggctg | agcacaatgc | ctcgggggtg | ataagccgct | 3060 |
| accgcgtccg | ccatatgcct | gagtgagggc | gctaaggctg | ggccgctcaa | tggcttcgag | 3120 |
| aataaagcgg | gccaggtgat | cttctgccat | ctctccgaa | tgggtggaag | cacatcaacc | 3180 |
| 20tgaccagcga | ttacgtctgc | gcagagccgc | agactagagg | tcgggaagtt | ccggccgtta | 3240 |
| cggctgcccc | gaaaacctca | gtgtacgatt | ttttcagtt | cctctattcg | acctgcgttg | 3300 |
| cgacgttcga | tctgactagg | gcgcagcggc | aacatcggt | gagtacgatc | cagcgcttg | 3360 |
| atttgggtta | tttcgttaac | cgacaacacc | agagcggtat | cctgcgggtt | caggcagagc | 3420 |
| ccaaccacat | gactacctct | cggcaaaatg | agggtcggtta | ctaattttga | acgtcttcaa | 3480 |
| 25gcgatgggg | tttaggtcga | aagctgcccc | ggcctgcgga | acctaccaca | tcgtaaccgg | 3540 |
| agcatattcg | gcccgcggcg | gctaaaaaca | cagaaatgag | cgggggtgacc | cgatgcctt | 3600 |
| tgatcgattc | tccgctttca | aaagcggcag | gggctgaagt | cagccggaaa | taccagacc | 3660 |
| aatcacaaca | ttcatgctgg | tgataaataa | attcaactat | gctttattga | caaataaaag | 3720 |
| cacactcacc | atcatcgca | atacaaatct | tataaaatta | agccggattt | atgaaacttc | 3780 |
| 30tcatacagcc | aaacaatcgc | ctcattagct | ttagtcccg | cgccaaacct | ctggaagtgc | 3840 |
| ttcgcgagaa | cgggtgtcgt | atttcctaca | gttgtatgtc | tgggcgttgc | ggaacctgcc | 3900 |
| gggtgccggg | tacagatggg | agtgtaatgt | attcggggac | gggaagcggg | ttaccacacc | 3960 |
| tcgtggacga | gcattatgtg | ctcgctgtc | ggtcagtact | tactaacaat | tcgcgcatcg | 4020 |
| aaatcccaga | agccgacgaa | atcgtcacc | accggcgag | aatcatcaag | ggcactgtgg | 4080 |
| 35tcgctattga | gtcgccact | cacgatatcc | gtcgctacg | cgtacgcctc | gccaagccct | 4140 |
| tcgagttctc | accggacag | tacgcgacat | tgagttcag | tcctgagcat | gcgcgtccgt | 4200 |
| attcaatggc | aggctctgca | gatgaccaag | aaatggagtt | ccacatacgc | aagggtgccgg | 4260 |
| gtgggcgcgt | aactgagtat | gttttcgagc | acgtccgcga | aggtacaagc | atcaaatgga | 4320 |
| gcgggcccact | tggtacggct | tatttacgtc | agaaccacac | cgggcccgatg | ctctgtgtgg | 4380 |
| 40gcgggtggaac | cggactagca | ccgggtgctgt | cgattattcg | cggcgcgctg | aagttgggta | 4440 |

| | | | | | | |
|--------------|-------------|-------------|-------------|-------------|-------------|------|
| tgacaaaccc | catcctcctt | tatttcggag | tgcgagtc | gcaagacctc | tacgacgcag | 4500 |
| agcgattgca | taacctcgcg | gctgatcacc | ctcaactgac | cgtacacacg | gtaatcgcaa | 4560 |
| tgggccccgat | taatgagagt | cagcgagccg | gtctagttac | cgatgcgac | gaaaaagaca | 4620 |
| tcagttcgct | ggctgggtg | agggcctatc | tggtcggtgc | accagcgatg | gttgaagcgc | 4680 |
| 5tttgcaccgt | taccaaaccat | cttgggaatat | caccggaaca | tatttatgcc | gatgccttct | 4740 |
| atccccgtga | aatctgaatc | gtccctttcc | ctcacctcgg | tccattgagg | actcatcagg | 4800 |
| aggatactca | aataggcgca | aataataaca | gccgcgtcac | tatgtgcggc | agcgaaatgg | 4860 |
| tttccctctc | cctcatttgc | cccatcggag | gatagtttta | tgacagaaaa | atggattgaa | 4920 |
| gcagtcgccc | tttgtgacat | tccagaaggt | gatgtcctcg | gcgtgactgt | cgagggtaag | 4980 |
| 10gatctggcac | tgtacgaagt | ggaaggcgaa | atctacgcta | ccgacgacct | gtgcacgcac | 5040 |
| gggtccgccc | gcatgagcga | tggttatctc | gagggcgag | aaatcgaatg | ccccctgcat | 5100 |
| caaggtcggt | ttgacgtttg | tacaagcaga | gccctctgcg | cccccgtag | agagaacatc | 5160 |
| aaaacatatg | cagtcaagat | tgagaacctg | cgcgtaatga | ttgatttaag | ctgagaattt | 5220 |
| ttaataggcg | gcgccccgga | ccatagagcg | tgattatccc | cattccatct | tttttttaggt | 5280 |
| 15gaaaacatga | attacaaaaa | caaatcttg | gtaagtga | ctgggctgac | ccaaaagcac | 5340 |
| ctgattcatg | gcggtgaagg | gcttttccag | cacgaactga | gagccgtttt | tgcgcggaac | 5400 |
| tggttttttc | tactcatga | cagcctgatt | ccttcccccg | gcgactatgt | taccgcaaaa | 5460 |
| atggggattg | acgaggtcat | cgtctctcgg | caaagcgacg | gttcgattcg | tgcccttctg | 5520 |
| aacgtttgcc | ggcaccgcgg | caagacactg | gtgaacgcgg | aagccggcaa | tgctaaaggt | 5580 |
| 20ttcgtttgca | gttatcacgg | ctggggcttc | ggctccaacg | gcgaactgca | gagcgttcca | 5640 |
| ttcgaaaaag | agctgtacgg | cgagtcgctc | aacaaaaaat | gtctgggggt | gaaagaagtc | 5700 |
| gctcgctag | agagcttcca | tggttccatc | tatgcctgca | tcgatcagga | ggcccccttct | 5760 |
| cttatggact | atctcgggtga | cgctgcttgg | tacctggaac | ccatcttcaa | acattcaggc | 5820 |
| ggtttagaac | tggtaggccc | tccaggtaag | gttgatgatca | aggccaactg | gaaggcacct | 5880 |
| 25gcggaaaact | ttgtgggtga | tgcataccac | gtcggttggg | cgacgcgctc | ttcgctctgc | 5940 |
| acaggcgagt | ctatcttctc | gtcgctcgct | ggcaacgcag | tgctgccgcc | tgaagggtgcg | 6000 |
| ggcttgcaaa | tgacctcaa | atacggcagc | ggtatgggtg | tggttggtggg | cgatatttca | 6060 |
| ggcgtgcata | gcgcagactt | ggttccggaa | ttgatggcat | tcggcgggcg | taagcaggaa | 6120 |
| aggctgaaca | aagaaattgg | cgatgttcgc | gcccggattt | atcgagacca | cctcaactgc | 6180 |
| 30accgttttcc | cgaacaacag | cgtgctgacc | tgctcgggtg | ttttcaaagt | atggaacccg | 6240 |
| atcgacgcaa | acaccaccga | ggtctggacc | tacgccattg | tcgaaaaaga | catgcccgag | 6300 |
| gatctcaagc | gccgcttggc | cgacgcggtt | cagcgaacgg | tcgggcctgc | tggttcttgg | 6360 |
| gaaagcgacg | acaatgacaa | tatggaaaca | gcatcgcaaa | acggcaaaaa | atatcaatcc | 6420 |
| agagatagtg | atctgatttc | caaccttgg | ttcggggaag | atgtatacgg | cgacgcggtc | 6480 |
| 35tatcctggcg | tcgtcgaaaa | atcggcgac | ggcgagacca | gttatcgtgg | tttctaccgg | 6540 |
| gcttaccagg | cacacgtcag | cagctccaac | tggtctgagt | tcgaggatgc | ctctagtact | 6600 |
| tggcataaccg | aactgacgaa | gactactgat | cgtaacaga | cgagtcgacc | atgatgatca | 6660 |
| atactcaaga | agacaagctg | gtatccgccc | atgacgccga | agagtttctt | cgtttcttca | 6720 |
| attgccacga | ctcggttttg | caacaagaag | ccaccacgct | gctgaccg | gaagcgcac | 6780 |
| 40tgctggacat | tcaggcttac | cggacttgg | tagagcactg | cgtgggggtca | gaggttcaat | 6840 |

| | | | | | | | | | | | | | |
|----|-------------|----|------------|-----|------------|-----|-----------|-----|-----------|-----|-----------|-----|------|
| | atcaagtc | at | ttcacgcg | aa | ctgcgcgc | ccg | cttccgag | cgc | acgttata | ag | ctcaatga | ag | 6900 |
| | ccatgaac | gt | ttgcgacg | aa | aatttccag | c | aactgaa | agt | tcgagtcg | ag | catcaact | gg | 6960 |
| | attcacaaaa | | ctggagca | ac | agcccgaag | c | tgcgcttt | ac | tcgcttcac | t | accaatgt | cc | 7020 |
| | aggccgca | at | cgagctaa | at | gatgaaga | tc | tgcttcac | gt | ccgctcca | ac | gtcgttct | gc | 7080 |
| 5 | accggggcac | g | acgtggca | at | caagtcgat | g | tcttctac | gc | cggccggg | aa | gacaaatg | ga | 7140 |
| | aacgtggcga | | aggtggag | tg | cgaaaatt | gg | tgcagcgat | t | cgtggatt | ac | ccagagcg | ca | 7200 |
| | tacttcagac | | gcacaatc | tg | atgggtct | ttc | tgtgatcc | gg | tgaccact | ttt | tacaaatg | gt | 7260 |
| | gactgctacc | | gcggtcacca | | ttaatcaaaa | | gggaatgt | ac | gtgtatgg | gc | aatcaaca | ag | 7320 |
| | tcgtttcaat | | aaccgggt | gcc | ggctcagg | aa | tcgggtct | cg | actgggtc | cg | tccttcaag | t | 7380 |
| 10 | cggccgggtta | | ttgcgtat | cc | gctctcgt | ac | gaaacgag | ga | gcaagagg | cgc | cttctttg | ca | 7440 |
| | atgaattcaa | | ggacgcact | c | gagatcgt | tg | tgggcgat | gt | ccgagatc | ac | gcaataaat | g | 7500 |
| | agaagctgat | | caagcagac | a | atcgctag | at | tcgggtcat | ct | cgattgtt | tc | atcgcaaat | g | 7560 |
| | ccgggtatttg | | ggattacat | g | ctgagcat | cg | aagagcct | tg | ggagaaaat | c | tccagcag | tt | 7620 |
| | ttgacgaaat | | attcgacat | c | aatgtaaa | ga | gctatttc | ag | tggcatcag | t | gcagctct | gc | 7680 |
| 15 | cgggaactgaa | | aaagacga | ac | ggatcggg | tg | tgatgacc | gc | ttcgggtg | tcg | tcccatgc | gg | 7740 |
| | tcgggtggg | tg | tggttctt | gc | tacatcgcc | a | gcaagcat | gc | gggtgtag | gt | atggtcaag | g | 7800 |
| | ctttggccta | | cgaattgg | ct | cccgaatt | c | gcgtgaac | gc | tgtctcgc | cg | ggcggcac | cg | 7860 |
| | tgacgtctct | | gtgcggtc | ct | gcaagcgc | cg | gtttcgac | aa | aatgcacat | g | aaagacat | gc | 7920 |
| | ccggcatcga | | cgatatgat | c | aaaggcct | ca | ctcctctt | gg | gtttgcag | cc | aagcccga | ag | 7980 |
| 20 | acgtgggtgga | | gccctatc | tg | ttgctggc | tt | cgcgaaag | ca | gggaaaat | tc | atcacggg | ca | 8040 |
| | ccgtgattag | | cattgatgg | c | ggatggcg | c | tcgggtcg | caa | gtgagctt | gc | agccgatc | aa | 8100 |
| | aggttataga | | aacatttt | ta | ggtagcgcc | c | catgaaga | ca | aaactgtt | ta | tcaacaac | gc | 8160 |
| | ctggatcgat | | tccagtga | cc | agcagacc | tt | cgagcgca | ag | caccccg | tca | acagcgag | gt | 8220 |
| | gatgactgag | | agcgcaaac | g | ccacgggt | gac | ggacgcga | t | aaggcggc | gc | aagtggcc | ga | 8280 |
| 25 | ggagggcattc | | aagacctg | ga | aggacgtt | gg | accttcgg | ag | cgtcgcgc | gcc | ttctcctg | aa | 8340 |
| | ggtcgccgat | | gtcatggaa | a | gtaaaacac | c | caagtttat | c | gaagtgat | gg | ccatggag | gt | 8400 |
| | gggagcttcc | | gctctttg | gg | ccggattc | aa | cgtccatg | cg | tctgccaat | g | tgttccgg | ga | 8460 |
| | ggctgcctcg | | ctggccact | c | aaattcagg | g | cgaaaccat | c | ccaacgg | aca | aagccgaa | ac | 8520 |
| | gctctcaatg | | acactacgt | c | agccgggt | cgc | cccgatc | ctg | agcatcgt | cc | catggaac | cg | 8580 |
| 30 | caccgcagtg | | ctggcggc | ac | gagccatc | gc | gtatccgt | tg | gtctgtgg | ca | acacgggt | gt | 8640 |
| | gttcaaaggc | | tctgaatt | ta | gtcccgcg | ac | gcatgcct | tg | atcacccag | t | gcgtccag | ga | 8700 |
| | agccgggctg | | cccgtggc | g | tgctcaact | a | tctcaact | cc | tcgcctga | acc | gttcgccc | ga | 8760 |
| | gatcgccgac | | gcaatgat | ct | ccgcgaag | ga | gatccgcg | c | atcaactt | ca | cgggttca | ac | 8820 |
| | ccgcgtgggc | | agcattat | cg | cgcagaag | gc | cgcgaaac | ac | ctcaagcg | ct | gcctactg | ga | 8880 |
| 35 | gctcggcggc | | aagtcccc | gc | ttattgtt | ct | ggatgacg | ca | aacattga | cg | cggcgggt | caa | 8940 |
| | ggcagcggtg | | ttcggtag | ct | tcctgttc | ca | aggtcagat | c | tgcatgtc | ca | ctgagcgt | tt | 9000 |
| | gggtggtgat | | gagaagatt | g | ccgacgaat | t | tgctcgca | ag | tttgcgaa | a | aaactaag | cg | 9060 |
| | cttgagcgtg | | ggcgaccct | g | gcgttaact | gg | cgactgcac | t | atcggccca | a | tgggtctc | gcc | 9120 |
| | aaattcgggc | | gagcggat | ca | atgggttt | gtt | caaggacg | cgc | atcgataa | ag | gggccaa | agt | 9180 |
| 40 | tggttgccgc | | ggcatggccc | | aggggtcg | ggt | catgcggc | gcc | acgatcct | gg | atcacgtg | aa | 9240 |

| | | | | | | | |
|----|-------------|------------|-------------|-------------|------------|-------------|-------|
| | atctgacatg | cggatctacg | atgaggagac | ctttggtccc | atcacagtgg | tgatccggtg | 9300 |
| | caaaggggaa | gcagaggcca | tccgcattgc | caacgacagc | gtttatggtc | tgctcgctggg | 9360 |
| | cgtgttttggc | cgcgacatca | accgtgctct | gcgagtgggt | atgtcgatcg | aatatggttg | 9420 |
| | cgtacacatc | aacggctcga | ccgtccagaa | cgaggcgcag | gctccttacg | gaggtagcaa | 9480 |
| 5 | gaacaccggc | tacgggcgct | tcgacggccg | tgccgtggtc | gacgagttca | cagagctcaa | 9540 |
| | gtggctgacc | attgagccat | tcgagcagca | atatcccttc | tgaaagcact | aactccaagg | 9600 |
| | aatcaaacga | tgagtaagca | agctgcagtt | atcgagctcg | gatacatggg | catctcagtc | 9660 |
| | aaggatcctg | atgcgtggaa | atcgtttgcc | atgaatatgc | tgggtctgca | agtactcgat | 9720 |
| | gagggtgaga | aggaccgttt | ctatctgcgg | atggattact | ggcaccatcg | tatcgtagtt | 9780 |
| 10 | catcacggcg | gagaggacga | cttggaatat | ctaggctggc | gtgtagccgg | caagccggag | 9840 |
| | ttcgaagctt | tggggcaaaa | gctcattgat | gccggttaca | agatccgtgt | ctgcgcaaaa | 9900 |
| | gttgaggctc | aggagcgat | ggtgttgggc | ctgatgaaga | cagaagatcc | aggcggcaac | 9960 |
| | ccgaccgaga | tattctgggg | gccccggatc | gacatgagca | atccgttcca | tcccggccgc | 10020 |
| | cctctgcacg | gcaagtttgt | gaccggtgac | cagggccttag | gtcattgcat | cgttcgccaa | 10080 |
| 15 | accgacgtcg | ctgcggccca | taaattctac | agcctgctgg | gcttccgtgg | ggacgtcgaa | 10140 |
| | taccgcattc | cattgcccga | cggcatgact | gccgaactgt | cattcatgca | ttgcaacgcc | 10200 |
| | cgtgatcact | ccattgcggt | tggtgccatg | cctgctgcca | agcggtcaa | tcacctgatg | 10260 |
| | cttgagtaca | cccatatgga | agacttgggg | tatacgcatc | aacagttcgt | gaagaacgaa | 10320 |
| | attgacatcg | cettgcagct | tggcattcac | gccaacgaca | aggcgttgac | gttctacggc | 10380 |
| 20 | gcaacgcctt | cgggctggct | tatcgagccc | ggctggcgag | gtgctacggc | catagacgaa | 10440 |
| | gcggagtatt | acgtcggcga | catcttcggc | catggcgtcg | aggcacctgg | atatggcctg | 10500 |
| | gatgtaaaac | tgagttaaaa | gcgattatgc | gtacatcggg | tttctgacgt | tttctgcttt | 10560 |
| | ctatacagcg | caataacaat | aacaactaag | tggaatatatt | aataatgata | aaaaaacgt | 10620 |
| | ttctttacat | gcctcgcttt | acttggtcac | tcagcacccg | atgggccgaa | gaatcacctt | 10680 |
| 25 | ggacataaccg | catcggcatg | agtaatatg | cttttgatac | aagcgccaaa | gtgtacttgg | 10740 |
| | gaggccagcg | tgtgccggga | ggaagcgctg | acgcgagcaa | taacaacgcg | ctcacattcg | 10800 |
| | acttcggcta | tgccatcaac | gaccaatgga | atgcacgttt | gattgtcggt | attccacctt | 10860 |
| | caactaaagt | taagggcgca | ggcacacttc | cgggcattca | gctgggaaaa | ataacttatg | 10920 |
| | ctccaacatt | actgacgtta | aactataacc | tcccagcttt | tgggcccgtt | cgctctcaca | 10980 |
| 30 | tcgggtgcagg | agtcaattac | acgcgaattt | tggaaaagca | ggacgcta | ctaaaatcat | 11040 |
| | tcgatgccga | ccacgcttgg | tccccgcgc | tgcatgttgg | tgccgatatt | gacgtgagtc | 11100 |
| | gcaactggtt | cgtcagcatt | gatattcgga | agttatacct | gaaaaccgac | gcatcagggt | 11160 |
| | acttagggcc | acaggaggct | aaagcaaagg | taactcttga | cccattgata | acatcgatcg | 11220 |
| | caatcggacg | ccaattttga | tgcccccttt | taaggctctc | tatctatcta | actgcaaagg | 11280 |
| 35 | gtattttttat | gttgaataaa | attagtaaaa | ccgcgcgtct | taccgctgaa | gatatcaatg | 11340 |
| | gtgcctggac | tataatgccc | acaccgtcga | cgcctgatgc | ttctgattgg | cgcacgacta | 11400 |
| | acactgtgga | cttagacgag | actgcccga | tagttgaaga | gctgattgct | gctgggtgtca | 11460 |
| | acggtatttt | gagtatgggt | acctttgggtg | agtgcgccac | gttgacctgg | gaggagaaac | 11520 |
| | gtgattatgt | ttcgacggtt | gtcgagacca | ttcgcggtcg | tgtaccttat | ttctcgggca | 11580 |
| 40 | cgacggccct | gaatacccg | gaagtcattc | ggcagaccgc | agagcttata | gatattggcg | 11640 |

ctaacggcac catgctagge gtgccgatgt gggttaagat ggacctgccc acagcgggcc 11700
 agttctatcg tgatgttgca ggcgcggtac cggaggctgc cattgcgatt tacgccaaacc 11760
 ccgaagcatt caaattcgac ttccctcgcc cattttgggc agagatgtct aaaattcctc 11820
 aggtagtga tgccaagtat ctaggcacg gaatgcttga cttggacctg aaattggcgc 11880
 5ctaacatccg cttccttcca cagcaggacg actattacgc ggcgcacgc atcaatcccg 11940
 agcgcattac tgcgttctgg tcaagcggg ccatgtgcgg cccggctacc gctatcgct 12000
 tgcgtgatga agtggagcgg gccaaagagta ccggtgactg gatcaaggcc aaagccatct 12060
 ccgatgatat gcgtgcagcc gattcgacat tgtttccgcg tggcgacttt tcggagttct 12120
 cgaagtataa tatcgggctt gaaaaggcac ggatggacgc ggctgggttg ctcaaggctg 12180
 10gtccctgcg tctccctac aatcttgctc cagaagatta cctcgttggg gcacagaaat 12240
 caggcaagtc gtgggcccgc ctgcacgcta aatacagtaa agaattaaag tagttcacct 12300
 ccgcattgct gagcgcgagg ggtgacgtaa acgccgagcg gtgcgggaag taagtgaagt 12360
 agagttcatt tcttggtcca ggcactgcta gatcagcaaa gttagctgat ctagcagctc 12420
 cgaaaatttg ggcgaaagct gatcttagga atgcgggata aaggcagtac accgtaacga 12480
 15tcgggggtgt ccgttcatgt tgaacgacac cgctattgcg ccgacttctc ttcttcggag 12540
 tgtttgattg tgattgtoga tttctatttc gattttttga gtccgttctc ttacttgccc 12600
 aatcagcgtt tgtcaaagct tgcgcaagat catggcctta ccacgtgtta taacgcgatc 12660
 gatttgccgc gggcggaaat agcgatcggg aacgttggtc catctaactc cgacttgga 12720
 gtaaaatttg actatttgaa agtagatttg caacgggtggg cccagcttta cggaataccg 12780
 20ctgggtatttc cagctaacta caacagcaga cggatgaata ctgggcttta ttactcggag 12840
 gccgaggtgc aggcgcgtgc ctatgtgaat gtagtattta atgcgatttg gggagaaggc 12900
 atagcaccag atttggaag cttgcctgcc ctggtatctg aaaagctagg ctgggatcgt 12960
 agtgccttcg agcgtttct cagcagcaac gccgcaacag agaggtatga cgagcagaca 13020
 catttcgcca tcgagcga ggtgttcggg gtgccaacga tgtttttggg cgatgaaatg 13080
 25tggtggggga acgaccgtct gtttatgctc gagagcgcaa tggggcgctt gtgccgaaa 13140
 aatgccgatt taagtattg atctgatcgt tatttgctc atgagtcgct ttcaagatca 13200
 gcggatactg aagtcagcta aatgcgggac tacttcaggc catgcttcag gggcggtcag 13260
 ggcggactgg cgtgtgcacg gtagttgttg cgtgaggcat gtttttaaga tactattttc 13320
 agtgcctgct gctctctctg ggttgagtgt gagggcgccg atagtctcaa atctcgctt 13380
 30gtacctgacc agagatacgg gtgtttatga tggcatcgct ggtacagcg atgacgccat 13440
 gctgcttcac ttcgagtgt ttcacgacct cgacgctagc aatctcaacg agccagatga 13500
 tcgccagcca gggcggaagc tagtcgtgat tcgggttttc tacgttgtac ccgaatgcaa 13560
 tcttggtctt aacggtccac gccttaaagc tggaaatgcca aatctcgat gtcgacattg 13620
 gcccatataa tgttttggt ctcctctagt cagaaatcag taagggtgtg atggtgataa 13680
 35ctgggtcaagc cgggaatcag cggaatcgt gcgtggcg cactggcgat caaacagcg 13740
 tcaataacta ccattgaact ggcctaagag ggtgtggaca aaataaagta aacttcttgc 13800
 ctcttttcta tgcagattg agccatgcct aaaactggac gccaccgta atcgctacgc 13860
 accattaccc tctgctgacc aggcctcgct atacgcagcc ttattccagt caagccgaac 13920
 tggcgcaaga ttccatgcag aaacgggcat caccgcgcat cccgacacct ttgcaaaagc 13980
 40gctgagactg gccgggatcg ttcagataaa ggagcgtgcc aaaggtagct tccagcctcc 14040

cgagcctcgt aagtcttatg gctataccga ggcacaccgg cgtcagttgc cggagcaacg 14100
 ctatcccagt tgctgactg atgcagaatg gacgttggtc gctgatttat ttgaagtctc 14160
 gggaggtcgc ggcgtaccgc tcgccactcc cgacgcactc ttttggaagc ctgctgttat 14220
 gtcgtgcgca cagggtgctc ctggcgaatg ctgcttcgcg agtttcccta ttgggacaat 14280
 5gtctataaaa cctttcggcg ctggagcgtc agggcaagtt tgagcaaatg cacgaccggc 14340
 tccgtgccca atggcgagag cgggtggatc gtgatgaacg gccgtcagcc gctgttctgg 14400
 attcccaatc aacgcgcacg tctccccaag gcggtgaaag cggttacgac gcaggcaaaa 14460
 ag 14462

10

<210> 10

<211> 12808

<212> DNA

15 <213> Artificial Sequence

<220>

<223> A modified DNA sequence

20 <400> 10

gtcgacgcat gatgcaggac tgggccgac gcctcgacct cttcgagcag aaccaggctcg 60
 aggcggccag catgccgctc accgtgcac tggaagggtg gcccgcatc cggagtgcag 120
 aaaccgcaag cgcgccctct acgcgggttg ccgcttcgcc aagcctgctc gtgacgaagc 180
 cgggtgacgc catgccgttg gtttctgccc ccgcacatcg gctgccggca gtgccgcccc 240
 25agcgatcggc cgcgccgctg gtgccttcgg acattcagcg cgagagaatg gatttggttcg 300
 atgtcttcga agcgcgcac aaccttccc tcgctgcgtt tgccaagatg gcgggcaaat 360
 ctgcaggtg gatcagctac gagatccagg cgggcaactt gctggcggtg aacgtgggca 420
 accgcggcca gcgcgtgcc gactggcacc tcgacccgct caagcacgag ctgatccagt 480
 ctgtcctgaa gctgaccagg ggtgcggacc cttggcagat ctaccatgca ctgctgcagc 540
 30cgcgctcgat gctgcggggg cgctcggcac tggagggcgt gactgccagc aatctcgaca 600
 agctcgtcat ggccgtgagc acagcgggtg aggaaagcga atggaccccg ctgcgggtcg 660
 gtggtcgtgt agtcaaaca cgggccggta agggatggc tggacattat tcatattagt 720
 gatactaata ttcatattatg gtttattgac tattagaata cagtcaatga tcatggaacc 780
 ataaagcata taaataaaga agctagatct atggaaactt tcgtactacc gaacaatcgc 840
 35cgcttgccct ttgattccgg tgccaacctt ttggaagtgc tccgtgagca ccgtgtgggt 900
 atttcctaca gctgtatgtc tggacgatgc ggtacttgcc gctgccgagt tatagatggc 960
 agcgtcatta gttcggcggc gaaaagcggg gactcaaatc gcatcgaaga gcattatgta 1020
 ctgcctgtc agtcagtgt caccagcaat tgcgcaattg agatcataga ctgagacgac 1080
 atagtcactc acccggcgcg aatcatcaaa ggcatgggtg tcgccgtcga gtcgccact 1140
 40cacgatattc gccgatccg cattcgcctc gccaaagcct tcgagttctc acccggacag 1200

| | | | | | | |
|--------------|------------|-------------|-------------|-------------|------------|------|
| tacgcgatgc | tacagttcag | tcccgaacat | gtgcgtccat | attcaatggc | tggctcgcca | 1260 |
| gatgaccaag | aaatggagtt | ccatatccgc | aaagtgccgg | gcgggcgtgt | cacggagtat | 1320 |
| atcttcgagc | acgtcccgga | aggtacaagc | attaagttga | gtgggcctct | tggtagcgcc | 1380 |
| tatctgcgtc | aggctcacac | cgggccgatg | ctgtgtgtgg | gcggcgggac | cggactcgca | 1440 |
| 5ccggtgctgt | cgattgttcg | cggcgcgtg | aagtcgggaa | tgacgaaccc | catccacctt | 1500 |
| tatttcgggg | tgcgcagtc | gcaagacctt | tacgacgcag | accgattgaa | ccaactcgcg | 1560 |
| gctatccacc | ctcaactgac | tgtccataca | gtgatcgcg | cgggcccgat | taatgagggc | 1620 |
| cagcgggccc | gcctaattac | cgatttgatt | gaaaaagaca | ttccctcgct | ggctgggtgg | 1680 |
| agagcctacc | tgtgcggcgc | accagcgatg | gttgacgctc | tatgcaccgt | cgccaaagat | 1740 |
| 10cttggaatat | cgcccagaca | tatttatgcc | gacgccttct | atcccagcgg | ggtctgaatt | 1800 |
| gccccgaccc | ttcacctctg | tacatcgaga | attcatcagg | aagacactta | aatgagcatc | 1860 |
| actaacaaca | gccgcgtctg | aatatttagg | acagctggat | gatctctaac | tccatcatta | 1920 |
| ccccatttga | agatagcttt | atgacagaaa | aatggattga | cgcagtcgct | ctttatgaaa | 1980 |
| tccttgaagg | tgacgtcctc | ggcgtgacag | tcgaaggtaa | ggaactagcg | ctgtatgaag | 2040 |
| 15tggaaggcga | aatctacgct | accgacaacc | tgtgcacaca | tgggtgctgcc | cgcatgagt | 2100 |
| atggctttct | agaaggcaga | gaaattgaat | gtcctttgca | tcaaggtaga | tttgatgttt | 2160 |
| gcacaggcag | ggccttgctg | gcccctgtga | cacagaacat | caaaacatac | ccggtgaaga | 2220 |
| ttgagggcca | gcgtgtgatg | attgatttga | gctgagaatt | ttaataggag | gcaccccgga | 2280 |
| ccctagagcg | taatcccccc | cattcgatct | cttgagggtga | aaatatgaat | tacaaaaata | 2340 |
| 20aaaacttgg | gagtgaatct | gggctgaccc | aaaaacacct | gattcatggc | gacgaagaac | 2400 |
| ttttccagcg | cgaactggaa | accatttttg | ctcggaactg | gcttttcctg | actcatgaca | 2460 |
| gcttgattcc | gtcccctggc | gactatgtta | cggcaaaaat | gggggttgat | gaggttatcg | 2520 |
| tctccaggca | gaacgacggg | tcgattcgtg | cttttctgaa | cgtttgctgt | caccgtggca | 2580 |
| agacgctggg | acacgcagaa | gcaggtaatg | ctaaagggtt | cgtttgagc | tatcacggct | 2640 |
| 25ggggcttcg | cgctaaccgg | gaactgcaga | gcgtcccgtt | tgaaaaagaa | ctgtatggcg | 2700 |
| aggcgctcga | caagaaatgt | atgggattga | aagaagtcgc | tcgtgtagag | agcttccatg | 2760 |
| gcttcatcta | tggttgcttc | gatgaggaag | ccccttctct | caaagactac | atgggggacg | 2820 |
| ctggctggta | cctggagcct | atgtttaagc | attccggagg | gctagaactg | atcggtcctc | 2880 |
| caggaaagg | cataatcaag | gctaactgga | aagcgcgccg | ggaaaacttt | acgggggatg | 2940 |
| 30cgtaccacgt | gggttgagcg | catgcgtctt | cgcttcgctc | agggcagtcg | gtcttctcgt | 3000 |
| cgttagctgg | caacgcagct | ttgccccag | aagggtgcagg | tctgcaaatg | acctccaaat | 3060 |
| acggcagcgg | catgggtgtg | ttgtgggacg | gatattcagg | cgtgcacagc | gcagacctgg | 3120 |
| ttccggaatt | gatggccttc | ggcgtgtcta | agcaggaacg | gctgaacaaa | gaaattggcg | 3180 |
| aggttcgcgc | acgaatctat | cgcagccacc | tcaactgcac | cgttttcccg | aacaacagtt | 3240 |
| 35ttctgacctg | ctcgggtgtc | ttcaagggtat | ggcaccgat | cgacgcaa | accactgagg | 3300 |
| tatggacct | cgccatggtc | gaaaaagaca | tgcccagagga | tctcaagcgc | cgcttggtcg | 3360 |
| acgcggttca | gagaacggtt | gggcctgctg | gcttctggga | aagcgacgac | aacgacaata | 3420 |
| tggaacgggt | atcgcaaaac | gccaaagaa | atcagtcag | agatggcgat | ctggtttcca | 3480 |
| acctgggttt | cggcggggac | gtatacggcg | acgaggttta | tcctggcatc | gtcggcaaat | 3540 |
| 40cggcgattgg | cgagaccagt | tatcgtggct | tctatcgggc | ttacggcgcg | cacatcagca | 3600 |

gctctagctg ggctgaattc gaggatgtct ctaaaaattg gcataccgaa ctggcaaaga 3660
 ctactgatcg ctaacdgcg agagggacca tgatgattaa tattcaggaa gacaagcttg 3720
 tctccgcccc cgacgccgaa gagtttcttc gtttcttcaa ttcggcgac gaggctttgc 3780
 aacaagaagc taccacgttg ctaaccggg aagcgcatct tttagacatt caggcttacc 3840
 5gcgcctggtt agagcactgc gtggactcag aggtgaaata tcagattatc tcacgcgaac 3900
 tgcgctcagc ttccgagcgc cgttaccagc tcaatgaaac catgaacatt ttcaacgaga 3960
 attatgaaca actggaagtt cgcgtagcgc atcaactgga tccgcaaac tggggcaata 4020
 gtccaaaggt gcgctttact cgtttcatca caaatatcca ggctgcaatg gacgaaaatg 4080
 aagatttgct tcacattcgc tccaacctaa ttgttcaccg agcacgacgc ggcaatcaag 4140
 10tcgatgtctt ctatgccact cgggaggata aatggaagcg cggcgaagat ggagcgcgta 4200
 agttggtcca acgattgatt gattatccag agcgacatt ccagacgcac aatgtgatga 4260
 tctttatgtg acccaataat cgcctttaca aatggtgact gctacaagcg gtccatttgt 4320
 tcaaaaggaa atttatgtgt atgagcaatc aacaagtcgt ttcgataacc ggtgctggct 4380
 caggaattgg tctcgaactg gttcgatcct ttaagtccgc cggttattgc gtatccgctc 4440
 15tcgtacaaaa cgaggagcaa aaggcgagcc tttgcaatga gttcaaggac gactcgaga 4500
 tcgtcgtggg cgatgtccgg gaccacgcaa caaatgagaa gctgataaag caaacaaccg 4560
 atagattcgg ccattctgat tgtttcattg caaatgccgg tatttgggat tacatgcttg 4620
 gcacgaaga gccttgggag aaaatatcga gcagttttga tgagatatc aacatcaatg 4680
 tcaagagcta tttcagcggg atcaggggcg ccctgcagga actgaaaaag actagcggat 4740
 20cagtggatgat gaccgcttca gtgtcgtccc atgcggtcgg tgctgggtgt tcttgctaca 4800
 tcgccagcaa gcatgcgggc ctgggcatga tgaaagcttt ggcttacgaa ttggctcccc 4860
 acattcgcgt caacgccgta gcaccgggcg gactgtgac gcctctgagc ggtcccgcga 4920
 gcgccggtt cgacaaaact cacatggaaa acatgcccg tatcgaggac atgatcaagg 4980
 gtctaacgcc tcttggaaat tcagccaagg ccgaagacgt agtggcacc tatttgttgt 5040
 25tggcgctcgcg agatcaaggg aaattcatta cgggactgt cattaatata gatggagga 5100
 tggcgctcgg tcgcaagtag gtttgtcgcc tatcttga aaataactaa atttctggtg 5160
 aaaccgcgat aatacaaaat tgtttatcaa caatgtctgg atcaattcca gtgaccaaca 5220
 gaccttcgag cgaaagcacc ccgtcagtgg tgaggtgatg acggagtgtg caaactccac 5280
 ggtgatggat gcgttaaagg ccgcgcaagc tgcccaagag gctttccaga cctggaagac 5340
 30tggtggacct tcggagcgtc gccgccttct gctgagggtc gctgagggtt tggaaagtaa 5400
 aacaccgag tttatcgaag tgatggccaa ggaggtggga gcctccgctc tttgggccgg 5460
 cttcaatgtc cagatgtcag ccaatgtgtt ccgtgaagcg gcacgctgg ctacacaaat 5520
 tcagggggaa actattccga cagacaagtc tgacacgctc tcaatgacgc tacgtcagcc 5580
 ggtcgggtcc atctgagca tcgtgccgtg gaacggcacc gcagtgtgtg cggcacgagc 5640
 35catcgcttat ccgtggtct gcggcaacgc ggtggtattc aaaggttctg agtttagtcc 5700
 cgcgacgc at gccctgatca cccagtgcgt gcaggaagcc gggctgcctg ctggcgtgct 5760
 caactatctc aactcttcgc ctgaccgttc gcccgagatc gccgacgcac tgatctcagc 5820
 caaggagatc cgacgcac acttcacggg ctccaccgc gtgggcagta ttatcgcgca 5880
 gaaggccgcg caacacctca agcgtgcct gctggagctc ggtggcaagt cccacttat 5940
 40tggtctggat gatgcagaca tcgatgcggc ggtcaaggca gcggtgttcg gtagcttct 6000

gttccaaggt cagatctgca tgtccactga gcgcttgatc gttgatgaga agatagccga 6060
 cgaatttgtc gcaaaatttg tcgaaaaaac taagcgcttg agcgaggcg acccggtgcgt 6120
 aactggcgac tgcacatcg gcccgatggg ctcgccaaat tcgggtgagc ggatcaatgg 6180
 tttgttcaaa gacgcgatcg acaaaggggc aaaagttgtt tgcggcggtt tggcccaagg 6240
 5tgcgctcatg ccggccacga tcttgatca cgtcaaactc gacatgcgga tttacgatga 6300
 ggagaccttt ggtcccatca ccgtggtaac ccgttgtaaa ggcaagcag aggccgtccg 6360
 cattgccaac gacagcgtct atggcctgtc gtcgggcgta tttgggcgag acatcaaccg 6420
 cgctctacgc gtgggtatgt ccatcgata tggttctgta cacatcaacg gttcgaccgt 6480
 ccagaacgag gcgcaggctc cttacggagg caccaagaac accggctacg ggcgcttcga 6540
 10cgccgtgct gtaatcgagc agttcacaga gatcaagtgg ctgaccatcg aacctttcga 6600
 gcagcaatat cccttctgat aagcactaac tcccaggaat caaactatga gtaagcaagc 6660
 tgcagttatc gagctcggat acatgggtat ctcgggtcaag gacctgatg cgtggaaatc 6720
 atttgccacg gatatgctag gtctgcaagt tcttgatgag ggtgagaagg accgtttcta 6780
 tctgcggatg gattactggc atcatcgat cgtagtccat cacaacggac aggacgactt 6840
 15ggagtaccta ggctggcgtg tagccggcaa gccggagttc gaagctctgg gtcaaaagct 6900
 tattgatgcc gggtacaaga tccgcactcg cgacaaagtt gaggttcagg agcgtatggg 6960
 gttgggtctg atgaagacag aagatccggg cggaaccccg accgagatat tctggggccc 7020
 ccggatcgac atgagcaacc cgttccatcc cggtcgcccc ctgcacggaa agtttgtgac 7080
 cggtgaccaa ggcttggggc attgcatcgt tcgccaacc gacgtcgag aagctcataa 7140
 20gtttttatagc ctgctgggct tccgtgggga cgtcgaatac cggattccgt tgcccaacgg 7200
 catgactgcc gaactgtcgt tcatgcattg caacgcccgat gatcactcca ttgcttttgg 7260
 tgccatgccc gctgccaacc gactcaatca cttgatgctt gactacacc atatggaaga 7320
 cttgggatac acgcaccaac agtttgtaaa gaacgaaatt gacattgcct tgcagcttgg 7380
 cattcacgcc aacgacaagg cgttgacgtt ctatggtgca acgccttcgg gctggctcat 7440
 25tgagcccgcc tggcgagggt ccacggccat agatgaagcg gagtattacg tcggcgacat 7500
 cttcgcccat ggctggagg ccactggata tggcctggat gtaaaactga gctaaagatg 7560
 cgcgctcgtt gggcgagggt ctagtccagc atcttcatac gcaaccaacc ttgcagggcg 7620
 atgagatcaa aggacgttaa agcgaagggg aagtgggtcg ggccatgcgc ataccgatcc 7680
 atgacatttg tttcatagta tataggtaga taggtgaatc aagcgcttag tcaactagt 7740
 30gacacatctg tccatgagg ctatctacta tctattcaaa acaagaatga taaataggat 7800
 gaaaataata atgataaaaa gaacgatttg tcttggtgat cctctattct gtttggcaag 7860
 ccccatgagg gccgaagagt cgccttgagc gtaccgtatt ggtatgacta atgtagcttt 7920
 cgatgctagc gcaaaagtat acttaaatgg tcagcgggtg ccaggaggaa gcgctgatgc 7980
 gagegataac aacgcgctta cattcgactt cggctacgcc atcaacgacc agtggaatgt 8040
 35acgtgcgatt gtcggtatc cgctacaac taaagtgcg ggcgcaggca cacttcctgg 8100
 tatccagctg gggaaaaata cttacgctcc aacagtatta acgttgaaat ataacctccc 8160
 cgctttgggt cccgttcgcc ctcacatagg tgcgggagtc aattacacgc ggatttttga 8220
 aagtcgggac gctaactctaa aatcgctcga tgcgaccac gcttggtccc ccgcgctaca 8280
 tgttgggtgcc gatattgacg ttaaccgtgg ttggttcgtt agcattgata tccggaagtt 8340
 40atacctgaaa accgacgcat cagggtactt ggggccacag gaggtctaaag cacgggtaac 8400

| | | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|------------|------------|------|
| | tcttgacca | ttactaactt | cgatcgcgat | cggacgccaa | ttctgatgat | tctgttttaa | 8460 | |
| | gttctttatc | tatctaaccg | caaagggtgt | ttccatgtcg | aataaaatta | tgaaaacgtc | 8520 | |
| | gcgtcttacc | gccgaagata | tcaacggcgc | ctggactata | atgccacac | cctcgacgcc | 8580 | |
| | tgatgtcttc | gattggcgca | gcactgccac | cgtggactta | gaagagactg | cccgcatagt | 8640 | |
| 5 | tgaagagctg | attgcagctg | gtgtcaacgg | tattctaagt | atgggtactt | ttggtgagtg | 8700 | |
| | cggcacgttg | acctgggatg | aaaaacgtga | ttatgtctcg | acgattgtcg | agaccattcg | 8760 | |
| | tggtcgcgctg | ccttattttct | gtggcacgac | agccttaa | at | acccgagaag | tcattccgcc | 8820 |
| | gacccgagag | cttatcgata | ttggcgccaa | cggcactatg | ctcgggggtg | cgatgtgggt | 8880 | |
| | gaagatggac | ctgcctacag | cggttcagtt | ctatcgatg | gttgcatg | cggtagcaga | 8940 | |
| 10 | ggctgccatt | gcgatttacg | ccaaccccg | agcattcaag | ttcgacttcc | ctcgccatt | 9000 | |
| | ctgggcagag | atgtccaaaa | ttccgcaggt | agtgactgcg | aagtatctag | gcacgcgaat | 9060 | |
| | gcttgacttg | gacctgagac | tggcacccaa | catccgcttc | cttccccacg | aagatgacta | 9120 | |
| | ttacgcggcc | gcacgcac | atcccgagcg | cataaccgcg | ttctgggtcaa | gcggggccat | 9180 | |
| | gtgcggcccg | gctaccgcc | tcattgttgcg | tgacgaagtg | gtgcgggcca | agagcaccgg | 9240 | |
| 15 | tgactgggcc | aaggccaaag | ccatctccga | tgatatgctg | gcagccgact | cgacattggt | 9300 | |
| | tccgcgtggc | gacttttcg | agttctcgaa | gtataatatc | gggcttgaaa | aggcacggat | 9360 | |
| | ggatgcggct | ggttgggtca | aggctgggcc | ctgccgtccg | ccctacaacc | ttgttccaga | 9420 | |
| | agactacctc | gctggtgcac | agaaatcagg | caaggcttgg | gccgcgctgc | acgctaaata | 9480 | |
| | cagtaatgaa | ttgaagtagt | tcacctccgc | agacctgagt | gacaggggtg | cgcagacgct | 9540 | |
| 20 | gaggggtgcg | gaattaagtg | agctaaagca | catttcttgc | gccaggcatt | gccagatcag | 9600 | |
| | caaagtttgc | tgatctggca | gtttcaaaaa | tttgggcgaa | agctgatatc | aggaatacgg | 9660 | |
| | gataaaggca | gtgcaccata | acgacggggg | cgtgccattc | gtgatgaacg | attttgc | tat | 9720 |
| | tgtgccgact | tctgttcttg | gagtgtttga | ttgtgattgt | cgatttttat | ttcgattttt | 9780 | |
| | tgagtccgtt | ctcttacttg | gccaaaccatc | gtttgtcaaa | gcttgcgcaa | gactatggct | 9840 | |
| 25 | tttccattcg | ttattacgca | atcgatttgg | cgcgagttaa | aatagccatc | ggaaaacgtt | 9900 | |
| | gtccatctaa | tcgcgacctg | atagtcaagc | tggactat | tt | gaaagtagat | ttgcaacggt | 9960 |
| | gggccgagct | ttacgaaata | ccgttggtat | ttccagctaa | ctacaacagc | cgacggatga | 10020 | |
| | atactgggct | ttattactcg | ggagccatgg | cacagactgg | tgcttatgtg | aatgtagtat | 10080 | |
| | ttaatgcgg | ttggggagat | ggcatagctc | cagatttgg | aagcttgcc | gtctgtgtat | 10140 | |
| 30 | ctgaaaaact | aggetgggat | cgtagcgct | tcgaggactt | tatcagcagc | gatgccgcaa | 10200 | |
| | cagagaggta | tgacgagcag | acacatgccg | cgatcgaacg | caaagtgttc | ggtgtgccaa | 10260 | |
| | cgatgttttt | gggcgatgaa | atgtggtggg | gaaacgaccg | tctattttatg | ctcgagaacg | 10320 | |
| | cagtgggagg | tgcgcctgta | aatggagaat | agtcgctacg | gagcgcttgt | gccggctaaa | 10380 | |
| | tgccgatata | agtggttgac | ctgatcggtta | tttgctcgat | acagcgcttt | caaaatcagc | 10440 | |
| 35 | ggctactgaa | gtcagataaa | aatgcgggac | tacttcaggc | atcctgtgcg | acacaaaagt | 10500 | |
| | ttacctgtaa | ttgtccacct | attccgagtt | tggaatggta | gctgactcgc | tatgcgacca | 10560 | |
| | gcgatagcct | aacaagacat | gcattcactgg | taacgggtggg | gtgtgaagct | cctgcaacaa | 10620 | |
| | tgtagccctt | tgatgtgtgt | atttgctgcg | aggtgaagca | cagatgctcg | gagccgtacc | 10680 | |
| | ggcttgtggc | gctaggctgg | caagtatgag | caacgtaagt | gggggttggg | gcgcaatggg | 10740 | |
| 40 | aacacaaaaac | caacgcaagc | cttaccagcg | tcgttcggtg | ccttctctccc | atgcctccgc | 10800 | |

ctcgataaag cagctgcgca tateggcttc ctggctgac tcggttagta ggtcatgcaa 10860
 ggtcttgtcc agcgcctcgt cgctccgata cggaatggtc agctcgtaat ggccggtctc 10920
 cgaccgcttc atgccgtagg gctccaggca gtagcgctcg atgttctccg tggcccgctt 10980
 ccgaccgcgc atgaacttgc tgttggtcac caccgccagg cgcaggggta cggtgccac 11040
 5ccgctcgacg gttgactctg ccggtgacgc gatattgcgc ttttgacctc gcgccagggc 11100
 gctcttcttg tacgtcccga tctcgacgcc acggtggcgt aggtagctgt acaggggtgt 11160
 cttggagatg tgcaacttct cgccgatggc gctgacgctc aggcggccct cgcggtacag 11220
 ggtctccgcc gccatggcgg tggcctcggc cttggctggc aggccttgg gacggcgacc 11280
 gatccggcct cgagtccgtg ccgccgacag gccgcctga gtccgctcgc ggatcagctc 11340
 10gcgctcgaac tgaacagggt gaacaccagg cgatcttggg cgtgggtgct gtcaatgggg 11400
 tcgttcaggc tctgcaagcc gactttgctg gcagccagct agccgaccaa ctcaaccagg 11460
 tgcttgagcg agcgaccgag gcgatccagc ttccagatca ccacggcatc gcccgctcga 11520
 acatgggcta gcaacttgtc caactccggc cgcgcgcttt ttgcgccgct ggcgatgtct 11580
 tgatagatgc gttcgcaccc ggctgtttc agggcatcga cctggaggtc ggcgttctaa 11640
 15tcccgagtgc tcaccgcgt ataaccgatc ttcataaaaa gtaccgttta cttgactgca 11700
 ttagtaatag ttgaactttg attaatgcta ccagttattt gaaccctagc gcaggtgtaa 11760
 gcgtccagcc gcccacctc tactcagctt gatgaaccga ggggcagtag ttcacatc 11820
 cggctgttag ggaaactctg aagaagactt cctgattttg gcaaaatgcc cggatttcac 11880
 ccgccgagtt ttccaatgaa gcgatgacc ttcctcgacg ccgagtatgc cggtaagcgc 11940
 20aaacagaccc gcaaagagct gttcctgatc gagatggatc ggggtgggccc gtggaagggt 12000
 ttgattgcct tgatcgatcc gcattacccc aagggtgagg gtggcccgcc agcctatcca 12060
 ctgacggcaa tgctcgggt tcatctgatg cataagtggc tcggctacag cgaccggcg 12120
 atggaggaag cgctgtacga gacgaccatc ctgcgccagt ttgccgggtt gaggctggag 12180
 cgcattccgg actaaaccac catccccaac ttccgccggc tgctggagaa acacgaacta 12240
 25cctaccggca tcatgtttgc tgaattttcg cttttggctt tgagtgcggc attcaaccga 12300
 tacaggtgtg gcatgacacg ccaaagaaat caaccttggc ggctagaaca cttgtcgatg 12360
 aagcgaacaa tccgcaaacg ttaaggctac caggaaaaac ccaggatgcg gtagccgata 12420
 gcccgcgcg gcattgggtc cttccgctc aagagtgtca aaggatattg gactaatgca 12480
 gcatgcctcg accacatggc cgtggtccgt tcatctgttg gccatacgtc tatcgaatcg 12540
 30gacgcgggcg gcgacgacct cgggttgttg gcgttcggcc cagtcgatga gctcagtcag 12600
 ggacggcatc agcgaagtcc cgagcggggt gagcccgtag cgcacagatg ggggtgtcga 12660
 gggcgttacc tcgcgccaga tgagtccgtc gcgttcgagg tgacgcagcg tctcggtcag 12720
 catccggcgc gaaatgtcgg gcacggcgcg ggccagcgca ttgaatcgtt gcggaccctg 12780
 cgacaagggtg accaggatca gcgtcgac 12808

35

<210> 11

<211> 15024

40

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

5

<400> 11

| | |
|---|------|
| atcgtcctta agcgccgcac gcagccgttt aagtgaattg ctgacggcag gctgcgtcag | 60 |
| ccccagtttt tcgccggccg tcgatacgct ccggtcgagc agtagctggt tgaagaccac | 120 |
| cagcagattc aagtcgatgt cgcgagatc catgatgcct caccattatt catgctggtg | 180 |
| 10attttaacta tcagacttga tctatagcgc tataccgatc gacgcgccag aatcgcgacc | 240 |
| attcgagagac aactgaaaaa agagcttgca tggaaactggt agtagaacc cccaatttgc | 300 |
| atctgaacgc ggagaccggc agcaccctgc ttgacgtgct cagggtccaac gaggtcccca | 360 |
| tttcttatag ctgcatgtcg ggccgctcg gcaacttgccg ttgccgtgtg attgccggcc | 420 |
| atcttcgca taacggcccc gagacagggc gcccgaggc aggaaagga acctacgttc | 480 |
| 15tgccctgtca ggcggttctg accgaagact gcacgatcga gattcctgaa tctgacgaga | 540 |
| tcgtggttca cccggcgcg atcgtaagg ggacggtcac agcgatagac gaagccacc | 600 |
| atgacatccg gcgcctgctg atcaaactgg ccaaaccgct tgagttcagc cctggccagt | 660 |
| acgcaacggt gcagttcacg cccgaatgcg tccgccata ttcgatggcc gggctgccta | 720 |
| gcgatgcgga aatggagttt cagattcgcg cggttcggg cgggcatgtc agcaactacg | 780 |
| 20ttttcaatga actgtccgta ggcgcttcgg tgcggatcag cggccccctc ggaacggctt | 840 |
| atctgcccgc cacgcacacc ggccccatgc tttgtgtggg ggggtggaaca ggtctggcgc | 900 |
| ccgtcctttc gatcgcttca ggcgactgg aaagcgggat gagcaacccc atccatctgt | 960 |
| acttcggtgt gcggagcgag caggacatct atgacgagga acgccttcac gcattggctg | 1020 |
| caaggtttcc gaatctcaag gtgaatgtcg ttgttgcaac aggcctgcc ggccctggtc | 1080 |
| 25tgcatccgg cctggtcacc gatctgatcg gccgtgactt gcccaatttg gcgggatggc | 1140 |
| gcgcctacct gtgtggcgct ccggccatgg tcgaggccct gaacctgtc gttgctcgcc | 1200 |
| taggcatagt acccgggcac atccatgccg atgcgttcta tcccagcggc gtctgagcga | 1260 |
| aggcaccatg cgaacccaat tcaacccaag gataccaagc catgagtga cccaacgat | 1320 |
| taaaacccgt gtttcccaa gatccgaaat ggccgggga aggtagcagc cgcgttccct | 1380 |
| 30tctgggcta caccgcgaa gacctgtaca agcggaatt ggagcgctg ttctatgcaa | 1440 |
| accactggtg ctatgtaggc ctggaagccg agattccgaa tccaggcgac ttcaagcgaa | 1500 |
| cggatgatcg tgagcgctcg gtcacatggt tgcgtgatcc ggatggcggc atcaacgtgg | 1560 |
| tggagaacgt ctgcgccac cgtggcatgc gcttttgccg cgagcgccac ggcaacgcca | 1620 |
| aggacttctt ctgccctac caccagtggg actacagcct caagggtgac ctgcagggcg | 1680 |
| 35tgcccttcg ccgtggcgtc aagcaggacg gcaaggtaa cggcgcatg cccaaggact | 1740 |
| tcaaactcga agaacacggc ctgaccaagc tcaagggtgg cggccgaggc ggtgcagtgt | 1800 |
| ttgcctcttt tgaccacgat gtcgagccct tcgaggagtt cctggggcca accatcctgc | 1860 |
| attacttca ccgcgtcttc aacggccgca agctcaagat cctgggctac cgccgccagc | 1920 |
| gcatcccgga caactggaag ctgatgcagg agaacatcaa ggaccctac caccggggcc | 1980 |
| 40tgctgcacac ctggttctcg accttcgggc tctggcgcg cgacaacaag tcggaactga | 2040 |

| | | | | | | |
|--------------|-------------|-------------|-------------|-------------|-------------|------|
| agatggacgc | caagttccgc | cacgccgcaa | tgatctccac | gcgcggtcag | ggcggcaaga | 2100 |
| acgaggaggt | cgtgtccggc | gtggacagct | tcaaggaaca | gatgaagggtg | aacgacccgc | 2160 |
| gcctgtctga | catcgtgcc | gagccctggt | ggggcggtcc | gactgcgggtg | atgaccacga | 2220 |
| tcttccccag | cgtgatcatc | cagcagcagg | tcaacagcgt | atcgacccgc | cacatccagc | 2280 |
| 5ccaacgggtca | cggctccttc | gatttcgtct | ggacccactt | cggcttcgag | gacgacaacg | 2340 |
| aggagtggac | ccagcgccgc | ctgatccagg | ccaacctggt | cgggccggcg | ggcttcgtgt | 2400 |
| cggccgatga | cggcgagggtg | atcgagtgggt | cgcaggaagg | ctttgagcaa | aaaccgacgc | 2460 |
| accgcaccgt | gatcgagatg | ggcggtcacg | aaatcggcga | cacggaccac | atggtcaccg | 2520 |
| agacgtgat | ccgcgccatg | tacgactact | ggcgcaagggt | gatgggggaa | taaacatggt | 2580 |
| 10cgacttcaaa | acctatttcg | aactgctgaa | cctgtacagc | gactacgcc | tggtgtgcga | 2640 |
| ctccgccaat | tgggagaagt | ggcctgattt | cttcacgag | accggcacct | accgcctgca | 2700 |
| gccgcgcgag | aactttgagc | agggcctgcc | gctgtgcctg | ctggcgctgg | agagcaaggc | 2760 |
| aatgatccgc | gaccgggtgt | acggcgtcaa | ggaaccatg | taccacgacc | cctactacca | 2820 |
| gcgccacatc | gtgggcacgc | cgcgcgtgct | gtcagtggag | cgtgacgcgg | atggcgagcg | 2880 |
| 15catcaccgcc | gaagccagct | atgccgtgat | tcgcaccaag | tacgacggcg | attccacgat | 2940 |
| tttcaacgcc | ggctattacc | gagacgtgat | cgtgcgcacg | cccgaaggcc | tcaagctgaa | 3000 |
| gtcgcgcctg | tgcgtttacg | acagcgaaat | gatccccaac | tctgtgatct | atccaatctg | 3060 |
| aggcgacgca | catgactcag | aactggattg | atgcagcttg | tcttgacgac | atccctgaag | 3120 |
| gcgatgtggt | cggcgtcaaa | gttaacggca | aggaaattgc | gctctacgag | gtcgaggggg | 3180 |
| 20agatttatgc | caccgataac | ctgtgcacgc | acggcgctgc | gcgcatgagc | gatggctttc | 3240 |
| tcgaaggccg | ggaaattgaa | tgccctctgc | accaaggctc | atttgatgtc | tgcacaggca | 3300 |
| aagccttggtg | cacccccctg | acaaaggaca | tcaaaaccta | ccccgtcaaa | attgaaaaca | 3360 |
| tgcgcgtgat | gctcaaaatg | gagtaagact | gctttcaatc | aggaggtga | attcggaacca | 3420 |
| atcaacccaaa | ccatccaatc | caccatttag | ccaaaaaacg | gagacattaa | tcatgattta | 3480 |
| 25tgaaaatttg | gtgagtgaag | cagggctgac | gcaaaagcac | ctgattcatg | gcgacaaaaga | 3540 |
| acttttccag | cacgaattga | agaccatctt | cgcgcgaaac | tggttttttc | tgacccatga | 3600 |
| cagcttgatt | ccctcccccg | gcgactatgt | cacagccaaa | atgggtgtcg | atgaagtcat | 3660 |
| cgtctccgc | cagaacgatg | gctcgggtgc | agcctttttg | aatgtttgcc | gtcaccgggg | 3720 |
| caagacacta | gttcacgctg | aagccggaaa | tgcgaaaggc | tttgtgtgca | gttaccacgg | 3780 |
| 30ctggggcttt | ggctccaacg | gcgaactgca | aagcgttccc | tttgaaaaag | agttgtacgg | 3840 |
| cgatacgatc | aaaaagaagt | gcctgggctt | gaaagaagtc | ccccgcacgc | aaagctttca | 3900 |
| tggtttcatt | tacggttggt | ttgatgcaga | ggccccacg | cttgctcgatt | atctgggtga | 3960 |
| tgcagcctgg | tacttagaac | ccatcttcaa | gcactctggt | ggcctggaac | ttgtaggccc | 4020 |
| ccccggcaaa | gtggtgatca | aggccaactg | gaaggctcct | gcggaaaact | ttgtgggtga | 4080 |
| 35cgcgtagcac | gttggttga | cgcacgcac | gtctttgcgc | tcagggtcagt | cgatatattac | 4140 |
| ccctcttgcg | ggcaacgcta | tgcttccacc | cgaaggcgcg | ggcttacaaa | tgaccagcaa | 4200 |
| gtatggcagt | ggaatgggcg | tattgtggga | cggctactcc | ggtgtccaca | gtgctgacct | 4260 |
| ggttccccgaa | atgatggcat | tcggcggcgc | aaaacaggaa | aaactcgcca | aggaaatcgg | 4320 |
| cgatgtccgg | gcacggattt | accgcagcca | tctaaactgc | acggttttcc | cgaacaacag | 4380 |
| 40cattttgacc | tgctccgggtg | tcttcaagggt | ctggaacccg | atcgatgaaa | acacgaccga | 4440 |

ggtttggacg tatgccatcg tagaaaaaga catgcctgag gacttaaagc gtcgcttggc 4500
 tgacgcggtt cagcgcactg tcggaccagc aggattctgg gaaagcgacg acaacgacaa 4560
 catggagacg gagtcgcaaa atgccaagaa ataccaatcc agcaacagtg atctgattgc 4620
 caatttgggt ttcggaagg acgtctacgg cgacgaatgc tatccgggcg tcgttgccaa 4680
 5atcggcaatc ggcgaaacca gctatcgagg attctaccgt gcctaccagg ctccacatcag 4740
 cagctccaat tgggccgagt tcgaaaacac ctcccgaaat tggcacaccg aactcaccaa 4800
 gacgactgat cgctaatacca ggagccaacc atgatgatca ataccagga agacaagctg 4860
 gtctccgcgc acgacgccga agaatttcat cgtttcttcg tcgggcacga cagcgatctg 4920
 cagcaagaag tcaccacact cctgacccga gaagcacatc tgttggacat tcaggcctac 4980
 10aacgcctggc ttgaacactg cgttgcccc gagatcaaat accaagtgat ctgcgagaa 5040
 tttcggcca cttccgagcg tcgatacaca ctgaatgatg cggatgaacat ctacaacgag 5100
 aactatcagc acctgaaagt tcgagttgag catcagatgg acccgagaa ctgggccaac 5160
 agcccgaaga tccgtttcac ccgcttcgtc accaatgtca cagcggccaa ggacaagatc 5220
 gtaccggatc tgctgcatgt gcgttccaac ctcatctcc acccgccag acgcggaat 5280
 15caagttgacg tcttctatgc aacgcgggaa gacaaatgga aacgcacga aggtggtggc 5340
 atccaactgg tggaacgtct tgtggactac ccggagcgca ttctccagac ccacaatctg 5400
 atgaccttc tgtgaacct ggggatgcct gccttgatgg cggatcactc tgattgtttt 5460
 aaacagaaat ttattgccat gaacatacag caagttattg ccattactgg cgcgggttca 5520
 ggcacgggc tggaactggt tcgatccttc aaagcagctg gctattgctg gtccgcactt 5580
 20gttcgcaacg aggaacaaga ggcgggcctt cgagtgat tcaaagatgc cattgagatc 5640
 gtagcgggcg atgtttgtga tcacgccacc aatgaaaagc tggatcaaca ggcggtcgcc 5700
 aggttcggac acctcgactg cttcatcgga aatgccggga tatgggatta catgctgggc 5760
 gtcgacgagc catgggagaa attatccggc agtttcgagg agatatttga tatcaacgtc 5820
 aaaagctatt tcagcggcat cagcgcggcc ttgccggaac tcaaaaaaac gaacggatcg 5880
 25gtcgtagtga cggcttcctt ttcttctat gcggccggcg gcggcggtc ttgctacatt 5940
 gccagcaagc atgcagtgtt gggatgtgac aaggcgttg cctacgaatt ggccccgcac 6000
 atccgggtca atggcgttg gccagtggt acggtcactt ctttggctgg gccggcaagc 6060
 gccggtttcg acaaaaccaa aatgaaagac atgcccgga tcgatgacat gatcaagggc 6120
 ctgacccctc tggggttcgc agcaaggccc gaggacgtgg tggcaccgta tcttttgcgt 6180
 30gcctcccgga aacaaggga gttcatcact ggcaccgtaa tcggcattga tggcgcatg 6240
 gcgctcggtc gaaagtgaat tttcaatcaa atcagatttt tcaaccccat tcccaggaga 6300
 caacccatga agacgaaatt gttcatcaac aacacctgga gcgttcgag tgacaaaaag 6360
 tcattcgatc gcaagcacc tgtcagtggc gaggtcgtga ccaatgcgc gaacgccacg 6420
 gtggacgatg cggatcaatgc ggctcgagcc gctcaagagg cgttcaagtc ctggaaggcc 6480
 35gtcggaccct cggagcggcg gcgccttctt ttgaagggtg cagacgtcat ggagagcaaa 6540
 acgcccagat tcacgaagt gatggccaag gaagtgggag cctccgcgct gtggggggg 6600
 ttcaacgtgc acctgtcggc caatgtattc cgggaagccg cctcactggc caccacaatt 6660
 caaggcgaac ccattccgac ggacaagcct gacacctgt caatgacgct gcgtcacgct 6720
 gtcggcccca tcttgagcat cgttccttgg aacggcaccg ccgtgctcgc ggcgcgggcc 6780
 40atcgcttatc cgctggtctg cggcaatacc gttgtgttca aaggctccga gttcagcccc 6840

| | | | | | | |
|--------------|-------------|------------|-------------|-------------|------------|------|
| ggtagcgacg | ctttgatcac | caagtgcctg | caggaggccg | acctgcctgc | tggcgtgctc | 6900 |
| aactatctga | actcctcccc | ggaccggctg | cccgatattg | cagatgcgct | gatttcgtct | 6960 |
| aaagagattc | gtcgcacaa | cttcacaggc | tccactcgcg | tggggcgcat | catcgccag | 7020 |
| aaatcgggcc | agcatctcaa | gcgtgcttg | ctggagttgg | gtggaaaagtc | cccgtgatc | 7080 |
| 5gttctggacg | acgtgacat | cgacgaggca | gtcaaggccg | cgggtgttcg | cagcttcctg | 7140 |
| ttccaaggcc | agatctgcat | gtccaccgaa | cgcctggtgg | tcgacgaaaa | gatcgcgga | 7200 |
| gaatttgctg | cgaagtctgt | cgagaaaacc | aagcagttga | gcgcaggcga | tccatgcgtc | 7260 |
| acaggggact | gcatcatcgg | cccgatggtg | tcgcccact | cgggtgaccg | aatcaatggt | 7320 |
| ctgttcaaag | atgccatcag | caagggcgcc | aaggctcgtg | gcggcggcag | cgccgaggg | 7380 |
| 10gcggctcatg | ccgccacgat | cttggaccac | gtgacagccg | acatgcagat | ctacgatgag | 7440 |
| gaaacctctg | gtcccatcac | tgtggttctc | cgggtgcaaga | gcgaagcggg | cgccatccgc | 7500 |
| attgccaatg | acagcgcta | cggcctgtca | tcgggcgtgt | ttggccggga | cgtgaaccgg | 7560 |
| gctctgcgcg | tgggcatggc | gatcgaatac | ggctcgggtc | atatcaacgg | ctccaccgta | 7620 |
| cagaacgagg | ctcaggcgcc | ttatggcgga | acaaaggcca | cgggttatgg | ccgcttcgac | 7680 |
| 15ggacgcgcgg | tgatcgacga | gttcacggaa | ctcaagtggc | tgaccattga | accattcgag | 7740 |
| cagcagtatc | ccttctaagc | tgaagcaaca | aaggagttaa | accatgaaca | agccagcaac | 7800 |
| tgctattgaa | ttggggtaca | tgggcatctc | ggtcaaggat | cccgcagcgt | ggaaatcctt | 7860 |
| tgccgcaaac | atgctgggac | ttcaagtcct | cgatgagggt | gacaaggatc | gcttctatct | 7920 |
| gcgaatggac | aattggcacc | atcggatcgt | ggttcacac | aacgggtcaag | atgaccttga | 7980 |
| 20ataacctggc | tggcgtgctg | ccggtcaacc | ggaattcgag | gcattgggtc | aaaagctcgt | 8040 |
| ggacgcaggc | tacaaagtcc | gcgtgtgcga | caaagccgaa | gcacaagaac | ggatgggtgt | 8100 |
| gggcctgatg | aagacagaag | atccgggggg | caaccgcacc | gagattttct | ggggaccccg | 8160 |
| gattgacctg | aacaaccctt | tccatcccg | tcgtcccttg | cacgggaaat | ttctaaccgg | 8220 |
| tgatcagggc | ctgggccaact | gcacgtgctg | tcagaacgat | gttgaagcgg | cacgtaagtt | 8280 |
| 25ctatagcttg | ctgggatttc | gtggagatgt | cgagtaccgc | cttcctttgc | ccaacggcat | 8340 |
| gacggctgag | ttgacgttca | tgcatgcaa | tgctcggcat | cattcoatcg | ctttcgggtg | 8400 |
| aatgcctgctg | gccaagcgcc | tcaatcatct | gatgattgaa | tacactcata | tcgaagattt | 8460 |
| gggttgacac | caccagcttt | tcacgaagga | aaagattgac | attgccttgc | aattgggcat | 8520 |
| ccattccaac | gataaggcgc | tgacgttcta | cggggcaaca | ccttcgggct | ggctgataga | 8580 |
| 30acctgggtgg | cgaggcgccc | ccgccattgc | tgaatcgga | tattacgtcg | gcgacatttt | 8640 |
| cggccacacc | atcgaggcca | ccggttatgg | attggacgtc | aaactgagct | agccatgtaa | 8700 |
| cagatgcgaa | atcgatcgca | tctgttttcc | ccattcagtt | ccatataaaa | aaaaggagac | 8760 |
| aatgatgatc | aaagaagcca | tttccctcgc | aggacttggg | atgctgatgc | tcagcaccgc | 8820 |
| atatgccgag | gattcccggt | ggcatatcg | catcgcgcc | accaacgtag | ccttcgatgc | 8880 |
| 35gagcgccaaa | gtttcgattg | acggaacaag | ggtgccgggt | ggaagcgtg | acgccagtga | 8940 |
| caacaacgct | ttgacatttg | acttcgggta | catcattaac | gataactgga | acgcgcgatt | 9000 |
| aattgttggc | attccacca | ccacaaaagt | gacgggcgca | ggcacgctgc | ctccgatttt | 9060 |
| gctgggcccgt | gtccaatatg | ctcctgcagt | tttgtcggcg | acctacaacc | tgccacagat | 9120 |
| gggattgggt | cgcccgtacg | tggggcggg | gatcaactac | actcgaattc | tgaaaagcaa | 9180 |
| 40agatgccaat | ctgacctcat | tcgatgcaga | tcattgcgtg | gcacttgtgc | tccacatcgg | 9240 |

| | | | | | | |
|--------------|-------------|-------------|-------------|-------------|-------------|-------|
| tgcagaagcg | aacatcaacc | gcgactgggt | cgtcagcttt | gatatccgaa | aactttatct | 9300 |
| aaaaacagat | gcategggg | ttcttgggg | tcaagttgct | acagcccgcg | taacgttgaa | 9360 |
| tccgctgctg | acgtcgattg | cgattggccg | gagattctga | tcggtccaca | ttgattccaa | 9420 |
| aattctgttt | gcatacccat | tttcaaagga | aatttgaatg | acaagaaaga | cgagcaaagc | 9480 |
| 5gggtgcgctg | accgccggcg | atattcaagg | cgcatgggtc | atcatgccga | ccccgtccac | 9540 |
| gccggatgcc | tcggactggc | gcagcacgca | cacggtcgat | ctcgacgaga | cggccccgat | 9600 |
| tgtcgaggag | ttgattgctg | ccggcgctca | tggcattctg | agccacggca | cctttggcga | 9660 |
| atgcgcgacg | ctgacgtggg | aggagaagcg | ggattttgtt | tcaacggctg | tggaaaccgc | 9720 |
| gcgcggtcga | gtgccctact | tctgcggcac | aacggccttg | aatacccgctg | aagtcatacg | 9780 |
| 10ccagacccgc | gaattgatcg | acattggcgc | ccaaggaaca | atgctcggcg | tgccgatgtg | 9840 |
| ggatgaagatg | gatctgccta | ctgccgtgca | attttatcgc | gatgtggcgg | aagcagtgcc | 9900 |
| agatgcagcc | atcgctgtct | acgccaaccc | ggaggctttc | aaatttgatt | ttcctcgccc | 9960 |
| gttttggggc | gaaatgtcca | aaatcccgca | ggttgctaca | gccaagtact | tgggcatcgg | 10020 |
| gatgctggac | ttggatctga | aattggcccc | aaatattcgc | ttccttcgcg | atgaggatga | 10080 |
| 15ctactacgct | gcggcccggg | tcaatcccga | gcgcgatgact | gctttctggt | ctagcgggtc | 10140 |
| catgtgcggc | ccagcgaccg | cccttgctgt | gcgcgatgag | gtggtaaagg | ccaaaaatac | 10200 |
| aggatgattg | gccaaaggcca | aggctatttc | agatgacatg | cgcgcagccg | atgccacact | 10260 |
| gtttccacgc | ggcgatttct | cggaattctc | aaaatacaac | attggcctcg | aaaaagcacg | 10320 |
| aatggacgag | gccggtggc | tcaaggcggg | gccgtgccgg | ccaccctata | cgtggttcc | 10380 |
| 20cgacgaatac | cttgacggtg | cccgaatac | aggcaaggcc | tgggccgcac | tgcatacca | 10440 |
| gtatgccaaag | gaattgagga | aaacccaaac | ggcaaccaac | tcgaaaaaga | agtaagtcca | 10500 |
| ggccctgagt | cagacatctc | cgatcagcac | aacctgctga | tctggagggt | ttctggatta | 10560 |
| gtgcagtcgg | cgagttaaaa | atatgccagt | acagagtggg | agcgcacaaa | gcgcagggtg | 10620 |
| caccgatcgc | cctgaagcaa | tctcacacta | ttgccatctt | ctctttgttg | ggagtgcagt | 10680 |
| 25atcgtgatgg | tcgattttta | tttcgatttt | ttgagcccat | tttcgtatct | ggccaaccac | 10740 |
| cgtttgcctg | tgctcgccgg | gcgttatgga | ttctccatcc | agtatcacgc | cattgatattg | 10800 |
| gcgcgagcaa | aaacggccat | tggcaacatc | ggccatcca | atcgggacct | caaggtaag | 10860 |
| cttgactact | taaagggtga | tttgcagcga | tgggccgac | tctataggat | tccgttggtt | 10920 |
| ttcccccta | acttcaacag | ccgccgggtg | aatgccggac | tgtattaccc | ggcagccagg | 10980 |
| 30gagcgagccg | ctgaatatgt | tcgccttggt | ttcgattcgg | cttgggggaa | agggtgggca | 11040 |
| ctggatgctg | atagcttgct | ggctgaggta | tcgcacaagc | taaactggga | tctcggtgaa | 11100 |
| tttgaagatt | ttttgaacag | cgaaaatgcc | gccaaggcat | acgacgaaga | gacgcaggcg | 11160 |
| gccattgacc | gaaagggttt | cggggttccc | accgtgtttt | gggatgatca | aatgtgggtg | 11220 |
| ggaaatgacc | gccttttcat | gcttgagagc | agggtgcaaa | aggaaacgca | accataaatt | 11280 |
| 35ctcagtgcag | cgatttttct | gattttttcaa | ccacgccttg | atttttaagga | gtctcatgaa | 11340 |
| gctttattac | agccccggcg | catgttcac | gtcgcctcac | atcatccttc | gtgaagggtg | 11400 |
| atttgacttt | cagctagaaa | agggtgatct | cggcaccaag | gtgactgaga | ctggtgttga | 11460 |
| ttacaagacc | gtcaatcccg | tgggcagtg | ccccgctttg | caaattggatg | atgggcagg | 11520 |
| gctcaccgaa | gggccggcca | tcgttcaata | cctcgtgac | cgtgtgccag | aaaaatgcct | 11580 |
| 40ggcaccggca | gctggctcgc | tggagcggtta | ccgactgatg | gaatggctga | atttcatctc | 11640 |

| | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| | caccgaattg | cataaaagct | tcggaagcgt | gttcagcccg | gtgtttccac | aggatgccaa | 11700 |
| | gccagtcac | aaggcccaat | tggaaagccg | tcttgcccat | accgagcaga | tgctcgggtga | 11760 |
| | caaggtttgg | gccatgggga | acgacttttc | cgtgggtggac | gcctacttgt | ttactgtgct | 11820 |
| | tggctggggg | gcttacgtga | atgtggacct | ttcgcccttg | cccggcctgc | aaggctacct | 11880 |
| 5 | caaccgcgtg | gctgaacgtc | cagcagtcgg | ggccacattg | tcggtgaa | gtttgatctg | 11940 |
| | atTTTTTTtag | agaaataaat | aaagcatggc | cagcctccat | gagcggccag | ccatgccgcg | 12000 |
| | gggcagccgt | gggcagcggc | ctcgatcaaa | tttcgatgca | cacagattac | gagccccctt | 12060 |
| | ttccgactgc | ggtcggattc | agaaattgcc | gatacttgct | ttggtctgat | cgaagcgtga | 12120 |
| | gggatatcgg | cttccccctat | ccacatgctt | gatgaagagg | agcgtatcac | catgagtcac | 12180 |
| 10 | gaacttgcc | gactggaaga | cctgccgcag | gactaccgag | acgaactcaa | acaacttaac | 12240 |
| | ctggtgccgc | tatggcccag | cctacgcgcg | gtgctgccgc | ccaatgtccc | gacccgccag | 12300 |
| | acgcagccga | cttactggtc | ctaccagacg | ctcaagccgc | tgctgctcaa | ggccggtgag | 12360 |
| | ctgaccccc | ttgagaaggc | cgagcgccgc | gtgctggtgc | tggccaaccc | cggccacggc | 12420 |
| | ctggagaaga | tgcaagccag | cgccgccata | tacctgggca | tgcagttgct | gctgccgggc | 12480 |
| 15 | gagtgggcgc | ccagccaccg | ccacaccccc | aacgcggtgc | gcatgatcgt | ggagggcgag | 12540 |
| | ggcgctaca | ccaccgtgga | tggcgagaag | tgccccatga | gccggggtga | cctcatcctc | 12600 |
| | acgcccaccg | gcctgtggca | cgagcacggc | cacgacggca | acgagcccg | ggtgtggctc | 12660 |
| | gatgtgctgg | acctgccgct | ggtgtactac | atggaggcca | gctaccacat | cgacggcgag | 12720 |
| | cgccagcagg | tgcaccccg | cgggggcgac | tgcgcctgga | cccgctgcgg | cgtggtgcc | 12780 |
| 20 | accccggtgt | tccagcgcag | cgacaagcgc | tatcctctct | tgcgctaccc | ttggggccgac | 12840 |
| | acccgtgccg | ccctgctgtc | gctggcgggc | gaccagcctg | agcaggaatg | cgtgcaggct | 12900 |
| | acctacgtca | accccgaaac | gggtgacgac | gccgagaaca | tcttgggctt | ctacgccttg | 12960 |
| | atgctcaagc | ccggccagac | cctgcgcctg | cccgtgcgtt | cgcccgccgt | ggtgttccac | 13020 |
| | cagatcgaag | gccgcagcga | ggcgcgcatc | gccgagtcca | ccttcgccct | gagggaagcc | 13080 |
| 25 | gatacctgct | gcgcccccg | ctacaccgag | tgacgctga | aaaacctctc | agccgaccag | 13140 |
| | ccgtccttca | tcttcatggc | cgacgaatcg | cccctgcacc | gcaagctggg | cgtctttgaa | 13200 |
| | aaccgcggct | gagccggcgc | gaagcaccac | acaacaggaa | agcaagtgag | caactacctc | 13260 |
| | tggaaacccg | ctcccgctca | gtctctgccc | gttcgtggca | agaccgagcg | cttccccatc | 13320 |
| | aaccgcattc | tctgcgtcgg | ccgcaactac | cacgcccacg | cggtggaaat | gggccgctcc | 13380 |
| 30 | gtcgacaaaa | gcgtcgaaca | agcgtttctac | ttacacaaat | cgccacaaac | cctggtggaa | 13440 |
| | agtggtgcca | ccgtggccta | cccgcgcgcg | accagcaact | accactacga | gatggagctg | 13500 |
| | gtgctggcca | tggcgaagcc | cggcttccgc | gtcagcgaag | accaggcgca | cgagttgatt | 13560 |
| | tacggctacg | ccgccggcct | ggacatgacc | cggcgcgacc | tgcaactggg | ggcacgcgac | 13620 |
| | aaggggccgc | cctgggacac | cggcaaggac | atcgaggaag | gctcggctctg | ctccgagatc | 13680 |
| 35 | gtgccgatgc | aaggcgtggg | ggtggagcag | ggcgcgatcg | ccctggaagt | caacggccag | 13740 |
| | accaaacagt | cgtccaacgt | ggacaagctg | atctggaacg | tccgcgagat | cattgccgac | 13800 |
| | ctctccacct | actaccactt | gcaaccgggc | gacctgattt | acaccggcac | gcccgaaggc | 13860 |
| | gtggggcgccg | tgggtggcggg | tgacaagatc | atcgggccgtg | tggaaaggcat | tgccgagatc | 13920 |
| | agcctgaccg | tcgggccggc | cgagtgagcc | tgcacgcgat | gaagctgtac | aacttctggc | 13980 |
| 40 | qcaqcgccac | gtcacaccgg | ctgcgcacgc | cgctcaacct | caagggcgtg | ccctacgagt | 14040 |

acctggccgt gcacctgggc aaggaagagc acctaaagga cgccttcaag gcgctgaacc 14100
 cacagcagtt ggtgcccgcg ctggacacgg gcgcgcaagt gctgatccag tcgccggcca 14160
 tcatcgaatg gctggaggaa cagtaccca cgcgcgcgt gctgccggcc gacgccgacg 14220
 gcgcgcaacg ggtgcgcgcg ctggccgcca tcgtgggctg cgacatccac cccatcaaca 14280
 5accgcccgc tctggagtac ctgcgcaaga cgttcggcgc cgacgaggcc gccatcaacg 14340
 cctggtgtgg cacctggatc agcgcgggtt tcgatgccta cgaagcactg ctggcggctg 14400
 acccgaagcg cgcccgctac agctttggcg acacaccac gctggccgac tgctacctgg 14460
 tgccgcaggt ggaaagcgcc cgcgcgttcc aggtggacct gacgccctac cccctgatcc 14520
 gcgcagtgga cgcggcctgc ggcgagttgg acgcatttcg gcgcgccgcg ccagctgctc 14580
 10aacctgattc ggcttgacga aaaatacgcc gccttcgccc taataaacag ttcgtagccc 14640
 tgccagtgga caagtatcaa tgagcaatag tgaaaatgta aggggaaaga ttatgaaaca 14700
 tattctgacc cgacgagcag ccttgagcac tttgggtccc cttggaatcg ggggccttcc 14760
 gggcatgagc ttggcgcaat cgggccctgg cgtggccacg atcgtgattg gactcgccgc 14820
 agggcgagcg accgatatgg ctgccgcgcg gttgtctgag ggaatgcgcg gtgcctatgc 14880
 15ttccagtgtg ctgctcgaca accgcacggg ggcgggtggt cgaattgcga tccagcatgt 14940
 taaggccgcg gcaccaaacg gcgcgacact tctgctgacg cctgcatcca tgatgacgct 15000
 ataccctcac acgtacaaag atct 15024

20

<210> 12
 <211> 4912
 <212> DNA
 <213> Artificial Sequence

25

<220>
 <223> A modified DNA sequence

<400> 12

30 gagctcgttg cgcaagcgt tccatcagtg ggggcatgaa gtacatctcg ccgatgtcgg 60
 tcattgccaa gttgaagggt gcgctgctgg caaatgggtc gaaagagtca cgggtcgtca 120
 gtgcgctctg cagcgtggtg agcgcataga ccacgggctc cgcaagatgc agtgcatcacg 180
 gtgtcggctc catgcctttt gaggtgcgca agaacaaatc gtccttttagc gccgcacgca 240
 gccgtttaag tgaattgctg acggcaggct gcgtcagccc cagtttttcg ccggccgctg 300
 35 atacgctccg gtcgagcagt agctgggtga agaccaccag cagattcaag tcgatgtcgc 360
 gcagatccat gatgcctcac cattattcat gctggtgatt ttaactatca gacttgatct 420
 atagcgtat accgatcgac gcgccagtat cgcagccatt cggagacaac tgaaaaaaga 480
 gcttgcatgg aactggtagt agaaccctc aatttgcatc tgaacgcgga gaccggcagc 540
 accctgcttg acgtgctcag gtccaacgag gtccccattt cttatagctg catgtcgggc 600
 40 cgctgcggca cttgccgttg ccgtgtgatt gccggccatc ttcgcgataa cggctccgag 660

| | | | | | | | |
|----|------------|------------|------------|------------|-------------|-------------|------|
| | acagggcgcc | cgcaggcagg | aaagggggcc | tatgtcctg9 | cctgtcaggc | ggttctgacc | 720 |
| | gaagactgca | cgatecagat | tcctgaatct | gacgagatcg | tggttcaccc | ggcgcgcatc | 780 |
| | gtcaagggga | cggtcacagc | gatagacgaa | gccacccatg | acatccggcg | cctgcgcatc | 840 |
| | aaactggcca | aaccgcctga | gttcagccct | ggccagtagc | caacgggtgca | gttcacgccc | 900 |
| 5 | gaatgcgtcc | gcccctattc | gatggccggg | ctgcctagcg | atgcggaaat | ggagtttcag | 960 |
| | attcgcgcg | ttccggggcg | gcatgtcagc | aactacgttt | tcaatgaact | gtccgtaggc | 1020 |
| | gcttcgggtg | ggatcagcgg | ccccctcgga | acggcctatc | tgcggcgcac | gcacaccggc | 1080 |
| | cccatgcttt | gtgtgggggg | tggaaçaggt | ctggcgcccc | tcctttcgat | cgttcgaggc | 1140 |
| | gcactggaaa | gcgggatgag | caaccccatc | catctgtact | tcggtgtgcg | gagcgagcag | 1200 |
| 10 | gacatctatg | acgaggaacg | ccttcacgca | ttggctgcaa | ggtttcgaa | tctcaaggtg | 1260 |
| | aatgtcggtt | ttgcaacagg | ccctgcccgg | cctggtcatc | gatccggcct | ggtcacccgat | 1320 |
| | ctgatcgggc | gtgacttgcc | caatttgggc | ggatggcgcg | cctacctgtg | tggcgctccg | 1380 |
| | gccatggtcg | aggccctgaa | cctgctcggt | gctcgccatg | gcatagtacc | cgggcacatc | 1440 |
| | catgccgatg | cgttctatcc | cagcggcgct | tgagcgaagg | caccatgcga | acccaattca | 1500 |
| 15 | acccaaggat | accaagccat | gagtgaaccc | caacgattaa | aacccgtggt | tccccaagat | 1560 |
| | ccgaaatggc | cgggcgaagg | tagcagccgc | gttcacctct | gggcctacac | ccgcgaagac | 1620 |
| | ctgtacaagc | gcgaattgga | gcgcctgttc | tatgcaaacc | actggtgcta | tgtaggcctg | 1680 |
| | gaagccgaga | ttccgaatcc | aggcgacttc | aagcgaacgg | tgatcgggtg | gcgctcggtc | 1740 |
| | atcatggtgc | gtgatccgga | tggcggcatc | aacgtggtgg | agaacgtctg | cgccactgt | 1800 |
| 20 | ggcatgcgct | tttgccgcga | gcgccacggc | aacgccaagg | acttcttctg | cccctaccac | 1860 |
| | cagtggaact | acagcctcaa | gggtgacctg | cagggcggtg | ccttcgcgcg | aggcgtcaag | 1920 |
| | caggacggca | aggtcaacgg | cggcatgccc | aaggacttca | aactcgaaga | acacggcctg | 1980 |
| | accaagctca | aggtggccgc | ccgaggcggt | gcagtgtttg | cctctttcga | ccacgatgtc | 2040 |
| | gagcctttcg | aggactgtgc | gtgtacgaca | gcgagatgat | ccccaactcc | ctcatctacc | 2100 |
| 25 | ccatttgagg | ttgcagaaca | tgtccgagaa | ctggattgat | gccatcgcac | gggacgctgt | 2160 |
| | gcctgagggc | gatgtgggtg | gagtcatcgt | ggcaggcaaa | gacattgcct | tctatgaggt | 2220 |
| | ggaaggtgag | gtcttcgcca | ccgacaactt | gtgtaccac | ggggctgcgc | gcttgagcga | 2280 |
| | cggctttctc | gaaggccggg | aaattgaatg | tcctttgcat | caaggccgat | tcgatgtttg | 2340 |
| | cacgggtaaa | gccttgtgca | cacccctgac | acaggacatc | aaaacctacc | ccgtaaaaat | 2400 |
| 30 | cgaaaaatg | cgcgtgatgc | tcaagctgga | ctaaaactct | ttgcaggagg | aaagccaaat | 2460 |
| | ccggaatca | ccccacccaa | cccaatcact | acccgttttc | aaacaagagg | agataagcaa | 2520 |
| | ttatgagtta | ccaaaactta | gtgagtgaag | cagggctgac | gcaaaagctc | ctgattcatg | 2580 |
| | gcgacaaaga | acttttccag | cacgaattga | agaccatctt | cgcgcggaac | tggctttttc | 2640 |
| | tgacccatga | cagtctgatt | ccctcccccg | gcgactatgt | cacagccaaa | atgggcgtcg | 2700 |
| 35 | atgaagtc | cgtctcccg | cagaacgatg | gctcggtgcg | agcctttttg | aatgtttgcc | 2760 |
| | gtcaccgggg | caagacacta | gttcacactg | aagccggaaa | tgcgaaaggc | tttgtgtgcg | 2820 |
| | gctaccacgg | ctggggctac | ggttccaacg | gcgaactgca | aagcgttccc | tttgaanaag | 2880 |
| | agttgtacgg | agatgcgatc | aaaaagaaat | gcctgggctt | gaaagaagtc | ccccgcacgc | 2940 |
| | aaagctttca | tggctttatc | tatggctggt | ttgatgcaga | agctcccccg | ctcatcgatt | 3000 |
| 40 | atctgggtga | tgcagcctgg | tacctggaac | ccaccttcaa | gcactctggt | ggcctggaac | 3060 |

ttgtaggccc ccccggaac gtggtggtta aggccaaactg gaagcctttt gcggaaaact 3120
 ttgtaggtga catctaccac gttggttga cgcacgcagc ggctttgcgc gcagggcagt 3180
 cggtatattag ttctcttcgc ggcaacgcta agcttccacc cgaaggcgcg ggcttgcaaa 3240
 tgaccagcaa gtatggcagt ggaatgggct taacgtggga ctactactcc ggtaacttca 3300
 5 gcgctgatat ggttcccgat ctgatggcat tcggcgccgc aaaacaggaa aaactcgcca 3360
 aggaaatcgg cgatgtccgg gcacggattt accgcagcat tctgaacggc acggttttcc 3420
 cgaacaacag ctttttgacc ggctccgcta ctttcaaggt ctggaaccgc atcgatgaaa 3480
 acacgaccga ggtttgacg tatgccttcg tagaaaaaga catgcctgag gacttaaagc 3540
 gtcgcttgcc tgacgcggct cagcgcagtg tcggaccagc aggattctgg gaaagcgacg 3600
 10 acaacgaaaa catggagacg ttgtcgcaaa atgccaaaga ataccaatcc agcaacagt 3660
 atcagattgc cagtttgggt ttcggcaagg acgtctacgg cgacgaatgc tatccgggcg 3720
 tcgttggaac atcggcaatc ggcgaaacca gctatcgcg attctaccgt gcctaccagg 3780
 ctcacatcag cagctccaat tgggcccagtc tcgaaaatgc ctcccgaat tggcacaccg 3840
 aactcaccaa gacaactgat cgctaattca ggagccaacc atgatgatca ataccagga 3900
 15 agacaagctg gtctccgcgc acgacgccga agaatttcac cgtttcttca tcgtacaaga 3960
 tgatgcacta ctgcaagaag tcaacacgct cctgaccgcg gaagcgcacc tgctggacat 4020
 tcaggcctac aaagcctggc ttgaacactg cgttgccccc gagatcaaat accaagtgat 4080
 ctccgcgaga cttcgtctca cttccgagcg tcgatacaca ctgaatgatg cgggtgaatct 4140
 ctacaacgag aactatcaac agctgaaagt tcgagttgaa caccagatgg atcctcagaa 4200
 20 ctggggccaac aacccgaaga tccgcttcac ccgcttcgct accaatgtca cggcggccaa 4260
 ggacaagagc gcaccggaaa tgctgcatgt gcggtccaac ctcatctctc atcgcgccag 4320
 acgagaaaaa caagttgacg tcttctatgc aacgcgtgaa gacaaatgga aacgcacga 4380
 aggtggtggt atcaaattgg tcgaacgctt tgtggactac ccggagcgca tccccagac 4440
 ccacaacctg ttggtcttcc tgtgaacctt ggggatgcct gcctggatgg cgggcattcg 4500
 25 tgattatttt taacagaaat ttattgccat gaacacacag caagttgttg ccatcactgg 4560
 cgccggctcg ggcattgggt tcgagttggt tcgctctttt aaggcagccg gttatcgct 4620
 atctgcactc gttcgcaacg aggagcaaga ggcgggtctt cgcagtgaat tcaaagacga 4680
 cattgagatc gtggcgggag atgtccgtga tcacgccacc aatgagaagc tggttaaaca 4740
 gacggttgcc aagttcgggc gcctggattg cttcatcgga aatgccggga tatgggatta 4800
 30 catgctgagc atcgatgatg cttgggagaa attctcgggc agtttcgacg agatatttga 4860
 catcaacgctc aaaagctatt tcagcggcat cagcgcggcc ttgccggagc tc 4912

<210> 13

<211> 6779

35

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

40

<400> 13

| | |
|---|------|
| atgcatacgt agcgggtggcg aaagaggcgc cgctggaaga atccgggtctg tagctctggc | 60 |
| agaagaccca aggcgagatc aaccgcaccg gactccatat cctccttcag attgccagca | 120 |
| ttcggggcgca gcggtctgat ctggatgtga ggagctcggt gcgcaagcgc ttccatcagt | 180 |
| 5gggggcatga agtacatctc gccgatgtcg gtcattgcca agttgaaggt gcgctgtctg | 240 |
| gcaaattgggt cgaaagagtc acgggtcgtc agtgccgtct gcagcgtgtt gagcgcatag | 300 |
| atcacgggct ccgcaagatg cagtgcatac ggtgtcggct ccatgccttt tgaggtgcgc | 360 |
| aagaacaaat cgctcgttcag cgctgtacgc agccgtttta gtgaattact gacggcaggc | 420 |
| tgcgtcagcc ccagtttttc gccggccgtc gatacgtctc ggtcgagcag tagctggttg | 480 |
| 10aagaccacca gcagattcaa gtcgatgtcg cgcagatcca tgacctcacc ctaccatta | 540 |
| ttcatgctgg tgattttaac tatcagactt gatctatagc gctataccga tcgacgcgct | 600 |
| agaatcgag ccattcggag acaactgaaa aaagagcttg catggaactg gtagtagaac | 660 |
| ccctcaattt gcatctgaac gcggagaccg gcagcaccct gcttgacgtg ctccaggtcca | 720 |
| acgaggtccc catttcttat agctgcatgt cgggcccgtg cggcacttgc cgttgccgtg | 780 |
| 15tgattgccgg ccatcttcgc gataacggcc ccgagacagg gcgcccgcag gcaggaaagg | 840 |
| gggcctatgt cctggcctgt caggcggttc tgaccgaaga ctgcacgac gagattcctg | 900 |
| aatctgacga gatcgtggtt caccggcgcc gcacgtcaa ggggacggtc acagcgatag | 960 |
| acgaagccac ccatgacatc cggcgccgtc gcatcaaact ggccaaaccg cttgagttca | 1020 |
| gccctggcca gtacgcaacg gtgcagtcca cgccgaatg cgtccgcccc tattegatgg | 1080 |
| 20ccgggctgcc tagcgatgcg gaaatggagt ttcagattcg cgcggttccg ggccggcatg | 1140 |
| tcagcaacta cgttttcaat gaactgtccg taggcgcttc ggtgcggatc agccggcccc | 1200 |
| tcggaacggc ctatctcgcg cgcacgcaca ccggcccat gctttgtgtg ggggggtgaa | 1260 |
| caggctctggc gcccgctctt tcgatcgttc gaggcgcact ggaaagcggg atgagctacc | 1320 |
| ccatccatct gtacttcggt gtgcggagcg agcaggacat ctatgacgag gaacgccttc | 1380 |
| 25acgcattggc tgcaaggttt ccgaatctca aggtgaatgt cgttggttga acaggccctg | 1440 |
| ccggccctgg tcatcgatcc ggccgtgtca ccgatctgat cggccgtgac ttgcccaatt | 1500 |
| tggcgggatg gcgactgcat cctgtgtggc gctccggcca tggtcgaggc cctgaacctg | 1560 |
| ctcgttgctc gcctaggcat agtaccggg ccatccatg ccgatgcgtt ctatcccagc | 1620 |
| ggcgtctgag cgaaggcacc atgcgaaccc aattcaaccc aaggatacca agccatgagt | 1680 |
| 30gaaccccaac gattaaaacc cgtgtttccc caagatccga aatggccggg cgaaggtagc | 1740 |
| agccgcgttc ccttctgggc ctacaccgc gaagacctgt acaagcgcga attggagcgc | 1800 |
| ctgttctatg caaaccactg gtgctatgta ggccgtggaag ccgagattcc gaatccaggc | 1860 |
| gacttcaagc gaacgggtgat cggtagcgc tcggtcatca tggtcggtga tccggatggc | 1920 |
| ggcatcaacg tggtagagaa cgtctgcgcc caccgtggca tgcgcttttg ccgcgagcgc | 1980 |
| 35cacggcaacg ccaaggactt cttctgcccc taccaccagt ggaactacag cctcaagggt | 2040 |
| gacctgcagg gcgtgccctt ccgccgtggc gtcaagcagg acggcaaggc caacggcggc | 2100 |
| atgcccaagg acttcaaact cgaagaacac ggccgtgacca agctcaaggc ggccgccccg | 2160 |
| aggcgggtgca gtgtttgcct ctttcgacca cgatgtcgag cctttcgagg acttctggg | 2220 |
| ccaaccatcc tgcactactt cgatcgcgtc ttcaatggcc gcaagctcaa gatcctgggc | 2280 |
| 40taccgcccgc agcgcacccc gggcaactgg aagctgatgc aggagaacat caaggacccc | 2340 |

| | | | | | | |
|--------------|-------------|-------------|-------------|-------------|------------|------|
| taccaccccg | gcctgctgca | cacctgggtc | tcgaccttcg | ggetctggcg | cgccgacaac | 2400 |
| aagtcggaac | tgaagatgga | cgccaagttc | cgccacgccc | caatgatctc | cacgcgcggt | 2460 |
| cagggcggca | agaacgagga | ggtcgtgtcc | ggcgtggaca | gcttcaagga | acagatgaag | 2520 |
| gtgaacgacc | cgcgctgct | cgacatcggt | cccagagccct | ggtggggcgg | tccgactgcg | 2580 |
| 5gtgatgacca | cgatcttccc | cagcgtgatc | atccagcagc | aggtcaacag | cgtatcgacc | 2640 |
| cgccacatcc | agcccaacgg | tcacggctcc | ttcgatttcg | tctggaccca | cttcggcttc | 2700 |
| gaggacgaca | acgaggagt | gacccagcgc | cgctgatcc | aggccaacct | gttcggggcg | 2760 |
| gcgggcttcg | tgctggccga | tgacggcgag | gtgatcgagt | ggtcgcagga | aggctttgag | 2820 |
| caaaaaccga | cgacccgcac | cgtgatcgag | atggggcggtc | acgaaatcgg | cgacacggac | 2880 |
| 10cacatggtea | ccgagacgct | gatccgcggc | atgtacgact | actggcgcaa | ggtgatgggg | 2940 |
| gaataaacat | ggtagacttc | aaaacctatt | tcgaactgct | gaacctgtac | agcgactacg | 3000 |
| ccatggtgtg | cgactccgcc | aattggggaga | agtggcctga | tttcttcac | gagaccggca | 3060 |
| cctaccgcct | gcaaccgcgc | gaaaacttcg | agcaggactt | gccgctgtgt | ctgctggcgc | 3120 |
| tggagagcaa | ggccatgatt | cgtgaccgag | tgtacgggtg | caaggaaacc | atgtaccacg | 3180 |
| 15atccctacta | ccagcgccac | atcgtaggca | cgccgcgcgt | gctgtcagtg | gagcgtgatg | 3240 |
| cgagcggcga | gcgcatcacc | gccgaagcca | gctatgccgt | gattcgcacc | aagtacgacg | 3300 |
| gcgattccac | gattttcaac | gccggctatt | accgagacgt | gatcgtgcgc | acgcccggag | 3360 |
| gcctcaagct | gaagtcgcgc | ctgtgcgtgt | acgacagcga | aatgattccc | aactccatca | 3420 |
| tctaccctat | ctgagaagga | atccaatgag | cgagaactgg | atcgacgccg | ccgcccgcga | 3480 |
| 20cgaggtgcca | cgagggcgac | gtgatcgcca | tcaatatcgt | cggcaaggag | attgcctcta | 3540 |
| cgaggtggcg | ggcgagatct | acgccaccga | caacacctgc | actcacggcg | ccgcccgcac | 3600 |
| gagcgatggc | tttctcgaag | gccgggaaat | tgaatgtcct | ttgcatcaag | gccgattcga | 3660 |
| tgtttgacag | ggtaaagcct | tgtgcacacc | cctgacacag | gacatcaaaa | cctaccccgt | 3720 |
| aaaaatcgaa | aacatgcgcg | tgatgctcaa | gctggactaa | atgctcaagc | tggactaaaa | 3780 |
| 25ctctttgcag | gaggaaagcc | aaatccggaa | atcacccac | ccaacccaat | cactacccgt | 3840 |
| tttcaaacaa | gatgagacaa | gcaattatga | gttaccaaaa | cttagtgagt | gaagcagggc | 3900 |
| tgacgcaaaa | gcacctgatt | tatggcgaca | aagaactttt | ccagcacgaa | ttgaagacca | 3960 |
| tcttcgcgcg | gaactggctt | tttctgaccc | atgacagtct | gattccctcc | cccggcgact | 4020 |
| atgtcaaagc | caaaatgggc | gtcgatgaag | tcacgtctc | ccgccagaac | gatggctcgg | 4080 |
| 30tgcgagcctt | tttgaatggt | tgccgtcacc | ggggcaagac | aatagttgac | gctgaagccg | 4140 |
| gaaatgcgaa | aggctttgtg | tgcggttacc | acggctgggg | ctatggctcc | aacggcgaac | 4200 |
| tgcaaagcgt | tccttttgaa | aaagagttgt | acggagatgc | gatcaaaaag | aaatgcctgg | 4260 |
| gcttgaaaga | agtccccgcg | atcgaaagct | ttcatggctt | tatctatggc | tgttttgatg | 4320 |
| cagaagctcc | cccgtcatc | gattatctgg | gtgatgtagc | ctggctacctg | gaacccacct | 4380 |
| 35tcaagcactc | tgggtggcctg | gaactttag | gccccccgc | caaagtgggtg | gttaagggca | 4440 |
| actggaaggt | ttttgcggaa | aacttttag | gtgacatcta | ccacattgggt | tggacgcacg | 4500 |
| catctatttt | gcgcgcaggg | caggcgatat | ttgctcctct | tgccgggcaac | gctatgcttc | 4560 |
| cacccgaagg | cacgggcttg | caagcgacca | ccaagtatgg | cagtggaatt | ggcgtatcgt | 4620 |
| tggacgccta | ctccggtgtc | cagagcgctg | atctggttcc | cgaaatgatg | gcattcggcg | 4680 |
| 40gcgcaaaa | ggaaaagctc | gccaaagaaa | tcggcgatgt | ccgggcgcgg | atttaccgca | 4740 |

gccaaagtga cggcacggtt ttcccgaaca actgcttttt gaccggcgcc ggtgtcttca 4800
 aggtctttta cccgatcgat gaaaacacga ccgaggcttg gacgtatgcc atcgtagaaa 4860
 aagacatgcc tgaggactta aagcgtcgct tggttgacgc ggctcagcgc tctgtcggac 4920
 cagcaggata ctgggaaagc gacgacaacg acaacatggt gttgtcgcaa aatgccaaaga 4980
 5aataccaatc cagcaacagt gatctgattg ccgatttggg ttctggcaag gacgtctacg 5040
 gcgacgaatg ctatccgggc gtcgttagca aatcggcatt cagcgaaacc aaccatcgcg 5100
 gattctaccg tgcctaccag gtcacatca gcagctcaa ttgggcccag ttcgaaaaca 5160
 cctcccgaaa ttggcacacc gaactcacca agacgactga tcgctaatac aggagccaat 5220
 catgatgatc aatacccgag aagacaagct ggtctccgcg cagcagccg aagaatttca 5280
 10ccgtttcttc gtcgggcacg acagcgatct gcagcaagaa gtcaccacac tcctgacccg 5340
 cgaagccgac ctgctggaca ttcaggccta caaagcctgg cttgaacact gcgttgcccc 5400
 cgagatcaaa taccaagtga tctcgcgaga acttcgctcc acttccgagc gtcgatacca 5460
 actgaatgat gcggtgaata tctacaacga gaactatcaa cagctgaaag ttcgagttga 5520
 acaccagatg gatcctcaga actggtacaa cagcccgaag atccgcttca cccgcttcgt 5580
 15caccaatgtc acggcggcca aggacaagag cgcaccggaa atgctgcatg tgcgggtcaa 5640
 cctcattctc catcgcgcca gacgaggaaa ccaagttgac gtcttctatg caacgcgaga 5700
 agacaaatgg aaacgcatcg aaggtgggtg catcaaattg gtcgaacgct ttgtggacta 5760
 cccggagcgc agtccccaaa cccacaacct gatgatcttc ctgtgagccc tggggatgcc 5820
 tgcctggatg gcgggcatte gtgattatit ttaacagggc ggattcaaaa gtgaagtgca 5880
 20acaccctgga ttctcagtga tgagagtgga gtgctgcggg ctattcacia gcagttcgcg 5940
 gtagaccgcc agcggtgatc gtactcccag ccctttccct ggccgggttg tgatttcac 6000
 ggcaatcgca tcgagctgct cctggctgta gatgctcaga tccgtccctt tgggcaggta 6060
 ctggcgcacc aggccattca tgttctcggt ggagcctctt tgccaagggc tgtgcgggtc 6120
 gcagaagtac acggctatgc cggtttctc actgagcttc ttgtgcatgg acatctccc 6180
 25gccctggctg tacgtcatgc tcaaccgcat cggtgcgca atgccagca gcttgtccgt 6240
 gaaggcctgc aagacgttgg cagcactggc cggttgaaac tcaggcagct tgaccagcat 6300
 caccagccgg ctggtgcgtt ccaccaaggt cccgactgcg ctggcattgc cttctccctt 6360
 gatgagatca cttcccagt gcccgggaaa ctggcggtct tcgatctcgg gcgggcgtac 6420
 atggatgctc agcatgctcg gaatctggcc tctgcggtcc ttgcccttgc tgcgcggcag 6480
 30ccgcttggtg tgggcgtgac gcagggtggc gatcagctct ttgcgcagct cttccacggg 6540
 catggcatag atgcagttgt agatggtttc gtgtgacacg cggatttcat ggcccagggg 6600
 atacaaacgt gccagtgtca gggcaatctg ctccggcgac cagcgcagcc gcagcagatg 6660
 gatcaccagg caggacaaga tcgactcggg atgcagcttg atagctggtc tgccacagcg 6720
 tctgcgctgc agggactggc actgggcatg tgcgctgcca tagccagatg aagatgcat 6779

35

<210> 14

<211> 449

40

<212> PRT

<213> Artificial Sequence

<220>

<223> A polypeptide encoded by SEQ ID NO:3

5

<400> 14

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Tyr | Asn | Asn | Lys | Ile | Leu | Val | Ser | Glu | Ser | Gly | Leu | Ser | Gln | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | | |
| Lys | His | Leu | Ile | His | Gly | Asp | Glu | Glu | Leu | Phe | Gln | His | Glu | Leu | Lys | |
| 10 | | | 20 | | | | | 25 | | | | | 30 | | | |
| Thr | Ile | Phe | Ala | Arg | Asn | Trp | Leu | Phe | Leu | Thr | His | Asp | Ser | Leu | Ile | |
| | | 35 | | | | | 40 | | | | | 45 | | | | |
| Pro | Ala | Pro | Gly | Asp | Tyr | Val | Thr | Ala | Lys | Met | Gly | Ile | Asp | Glu | Val | |
| | | 50 | | | | 55 | | | | | 60 | | | | | |
| 15 | Ile | Val | Ser | Arg | Gln | Asn | Asp | Gly | Ser | Ile | Arg | Ala | Phe | Leu | Asn | Val |
| | 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Cys | Arg | His | Arg | Gly | Lys | Thr | Leu | Val | Ser | Val | Glu | Ala | Gly | Asn | Ala | |
| | | | | 85 | | | | | 90 | | | | | 95 | | |
| Lys | Gly | Phe | Val | Cys | Ser | Tyr | His | Gly | Trp | Gly | Phe | Gly | Ser | Asn | Gly | |
| 20 | | | 100 | | | | | 105 | | | | | 110 | | | |
| Glu | Leu | Gln | Ser | Val | Pro | Phe | Glu | Lys | Asp | Leu | Tyr | Gly | Glu | Ser | Leu | |
| | | | 115 | | | | 120 | | | | | 125 | | | | |
| Asn | Lys | Lys | Cys | Leu | Gly | Leu | Lys | Glu | Val | Ala | Arg | Val | Glu | Ser | Phe | |
| | | 130 | | | | 135 | | | | | 140 | | | | | |
| 25 | His | Gly | Phe | Ile | Tyr | Gly | Cys | Phe | Asp | Gln | Glu | Ala | Pro | Pro | Leu | Met |
| | 145 | | | | | 150 | | | | 155 | | | | | 160 | |
| Asp | Tyr | Leu | Gly | Asp | Ala | Ala | Trp | Tyr | Leu | Glu | Pro | Met | Phe | Lys | His | |
| | | | | 165 | | | | | 170 | | | | | 175 | | |
| Ser | Gly | Gly | Leu | Glu | Leu | Val | Gly | Pro | Pro | Gly | Lys | Val | Val | Ile | Lys | |
| 30 | | | 180 | | | | | 185 | | | | | 190 | | | |
| Ala | Asn | Trp | Lys | Ala | Pro | Ala | Glu | Asn | Phe | Val | Gly | Asp | Ala | Tyr | His | |
| | | 195 | | | | 200 | | | | | 205 | | | | | |
| Val | Gly | Trp | Thr | His | Ala | Ser | Ser | Leu | Arg | Ser | Gly | Glu | Ser | Ile | Phe | |
| | 210 | | | | 215 | | | | | | 220 | | | | | |
| 35 | Ser | Ser | Leu | Ala | Gly | Asn | Ala | Ala | Leu | Pro | Pro | Glu | Gly | Ala | Gly | Leu |
| | 225 | | | | | 230 | | | | | 235 | | | | 240 | |
| Gln | Met | Thr | Ser | Lys | Tyr | Gly | Ser | Gly | Met | Gly | Val | Leu | Trp | Asp | Gly | |
| | | | | 245 | | | | 250 | | | | 255 | | | | |
| Tyr | Ser | Gly | Val | His | Ser | Ala | Asp | Leu | Val | Pro | Glu | Leu | Met | Ala | Phe | |
| 40 | | | | 260 | | | | 265 | | | | | 270 | | | |

Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 5 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
 10 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 355 360 365
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 370 375 380
 15 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 20 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 435 440 445
 Arg
 25
 <210> 15
 <211> 449
 <212> PRT
 <213> Artificial Sequence
 30
 <220>
 <223> A polypeptide encoded by SEQ ID NO:4
 <400> 15
 35 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1 5 10 15
 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
 20 25 30
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
 40 35 40 45

Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
 50 55 60
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
 65 70 75 80
 5 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
 85 90 95
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
 100 105 110
 Glu Leu Gln Ser Val Pro Phe Gly Lys Asp Leu Tyr Gly Glu Ser Leu
 10 115 120 125
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 130 135 140
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
 145 150 155 160
 15 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
 165 170 175
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
 180 185 190
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
 20 195 200 205
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
 210 215 220
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
 225 230 235 240
 25 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
 245 250 255
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 30 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 35 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 40 355 360 365

Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 370 375 380
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 5 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 10 435 440 445
 Arg

<210> 16
 15 <211> 449
 <212> PRT
 <213> Artificial Sequence
 <220>
 20 <223> A polypeptide encoded by SEQ ID NO:5

<400> 16
 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1 5 10 15
 25 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
 20 25 30
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
 35 40 45
 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
 30 50 55 60
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
 65 70 75 80
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
 85 90 95
 35 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
 100 105 110
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
 115 120 125
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 40 130 135 140

Arg

<210> 17
 <211> 449 .
 <212> PRT
 <213> Artificial Sequence

5

<220>
 <223> A polypeptide encoded by SEQ ID NO:6

<400> 17

```

10 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Phe Gly Leu Ser Gln
   1             5             10             15
   Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
      20             25             30
   Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
15      35             40             45
   Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
      50             55             60
   Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
      65             70             75             80
20 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
      85             90             95
   Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
      100            105            110
   Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
25      115            120            125
   Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
      130            135            140
   His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
      145            150            155            160
30 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
      165            170            175
   Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
      180            185            190
   Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
35      195            200            205
   Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
      210            215            220
   Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
      225            230            235            240
40 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly

```


245 250 255
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 5 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 10 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 15 355 360 365
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 370 375 380
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 20 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Arg Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 25 435 440 445
 Arg

<210> 18

30 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

35 <223> A polypeptide encoded by SEQ ID NO:7

<400> 18

Met Asn Tyr Lys Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Thr Gln
 1 5 10 15
 40 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Arg

| | | | |
|----|---|-----|-----|
| | 20 | 25 | 30 |
| | Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile | | |
| | 35 | 40 | 45 |
| | Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val | | |
| 5 | 50 | 55 | 60 |
| | Ile Val Ser Arg Gln Ser Asp Gly Ser Ile Arg Ala Phe Leu Asn Val | | |
| | 65 | 70 | 75 |
| | Cys Arg His Arg Gly Lys Thr Leu Val Asn Ala Glu Ala Gly Asn Ala | | |
| | 85 | 90 | 95 |
| 10 | Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly | | |
| | 100 | 105 | 110 |
| | Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu | | |
| | 115 | 120 | 125 |
| | Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe | | |
| 15 | 130 | 135 | 140 |
| | His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met | | |
| | 145 | 150 | 155 |
| | Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His | | |
| | 165 | 170 | 175 |
| 20 | Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys | | |
| | 180 | 185 | 190 |
| | Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His | | |
| | 195 | 200 | 205 |
| | Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe | | |
| 25 | 210 | 215 | 220 |
| | Ala Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu | | |
| | 225 | 230 | 235 |
| | Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly | | |
| | 245 | 250 | 255 |
| 30 | Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe | | |
| | 260 | 265 | 270 |
| | Gly Gly Ser Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg | | |
| | 275 | 280 | 285 |
| | Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn | | |
| 35 | 290 | 295 | 300 |
| | Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp | | |
| | 305 | 310 | 315 |
| | Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met | | |
| | 325 | 330 | 335 |
| 40 | Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val | | |

145 150 155 160
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
 165 170 175
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
 5 180 185 190
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
 195 200 205
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
 210 215 220
 10 Cys Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
 225 230 235 240
 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
 245 250 255
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 15 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 20 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Val
 25 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 355 360 365
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 370 375 380
 30 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 35 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 435 440 445
 Arg

40

<210> 17

<211> 449

<212> PRT

| | | | | | | | | | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | | | 85 | | | | | | 90 | | | | | 95 |
| | Lys | Gly | Phe | Val | Cys | Ser | Tyr | His | Gly | Trp | Gly | Phe | Gly | Ser | Asn | Gly |
| | | | | 100 | | | | | 105 | | | | | 110 | | |
| | Glu | Leu | Gln | Ser | Val | Pro | Phe | Glu | Lys | Glu | Leu | Tyr | Gly | Glu | Ser | Leu |
| 5 | | | 115 | | | | | 120 | | | | | 125 | | | |
| | Asn | Lys | Lys | Cys | Leu | Gly | Leu | Lys | Glu | Val | Ala | Arg | Val | Glu | Ser | Phe |
| | | | 130 | | | | | 135 | | | | | 140 | | | |
| | His | Gly | Phe | Ile | Tyr | Gly | Cys | Phe | Asp | Gln | Glu | Ala | Pro | Ser | Leu | Met |
| | 145 | | | | | 150 | | | | | 155 | | | | 160 | |
| 10 | Asp | Tyr | Leu | Gly | Asp | Ala | Ala | Trp | Tyr | Leu | Glu | Pro | Ile | Phe | Lys | His |
| | | | | 165 | | | | | | 170 | | | | 175 | | |
| | Ser | Gly | Gly | Leu | Glu | Leu | Val | Gly | Pro | Pro | Gly | Lys | Val | Val | Ile | Lys |
| | | | | 180 | | | | | 185 | | | | | 190 | | |
| | Ala | Asn | Trp | Lys | Ala | Pro | Ala | Glu | Asn | Phe | Val | Gly | Asp | Ala | Tyr | His |
| 15 | | | 195 | | | | | 200 | | | | | 205 | | | |
| | Val | Gly | Trp | Thr | His | Ala | Ser | Ser | Leu | Arg | Thr | Gly | Glu | Ser | Ile | Phe |
| | | | 210 | | | | | 215 | | | | | 220 | | | |
| | Ser | Ser | Leu | Ala | Gly | Asn | Ala | Val | Leu | Pro | Pro | Glu | Gly | Ala | Gly | Leu |
| | 225 | | | | | 230 | | | | | 235 | | | | 240 | |
| 20 | Gln | Met | Thr | Ser | Lys | Tyr | Gly | Ser | Gly | Met | Gly | Val | Leu | Trp | Asp | Gly |
| | | | | | | 245 | | | | 250 | | | | 255 | | |
| | Tyr | Ser | Gly | Val | His | Ser | Ala | Asp | Leu | Val | Pro | Glu | Leu | Met | Ala | Phe |
| | | | | 260 | | | | | 265 | | | | | 270 | | |
| | Gly | Gly | Ala | Lys | Gln | Glu | Arg | Leu | Asn | Lys | Glu | Ile | Gly | Asp | Val | Pro |
| 25 | | | 275 | | | | | 280 | | | | | 285 | | | |
| | Ala | Arg | Ile | Tyr | Arg | Ser | His | Leu | Asn | Cys | Thr | Val | Phe | Pro | Asn | Asn |
| | | | 290 | | | | | 295 | | | | | 300 | | | |
| | Ser | Val | Leu | Thr | Cys | Ser | Gly | Val | Phe | Lys | Val | Trp | Asn | Pro | Ile | Asp |
| | 305 | | | | | 310 | | | | | 315 | | | | 320 | |
| 30 | Ala | Asn | Thr | Thr | Glu | Val | Trp | Thr | Tyr | Ala | Ile | Val | Glu | Lys | Asp | Met |
| | | | | | | 325 | | | | 330 | | | | 335 | | |
| | Pro | Glu | Asp | Leu | Lys | Arg | Arg | Leu | Ala | Asp | Ala | Val | Gln | Arg | Thr | Val |
| | | | | 340 | | | | | 345 | | | | | 350 | | |
| | Gly | Pro | Ala | Gly | Phe | Trp | Glu | Ser | Asp | Asp | Asn | Asp | Asn | Met | Glu | Thr |
| 35 | | | 355 | | | | | 360 | | | | | 365 | | | |
| | Ala | Ser | Gln | Asn | Gly | Lys | Lys | Tyr | Gln | Ser | Arg | Asp | Ser | Asp | Leu | Ile |
| | | | 370 | | | | | 375 | | | | | 380 | | | |
| | Ser | Asn | Leu | Gly | Phe | Gly | Lys | Asp | Val | Tyr | Gly | Asp | Ala | Val | Tyr | Pro |
| | 385 | | | | | 390 | | | | | 395 | | | | 400 | |
| 40 | Gly | Val | Val | Gly | Lys | Ser | Ala | Ile | Gly | Glu | Thr | Ser | Tyr | Arg | Gly | Phe |

405 410 415
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 420 425 430
 Glu Asp Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 5 435 440 445
 Arg

<210> 20
 10 <211> 449
 <212> PRT
 <213> Artificial Sequence
 <220>
 15 <223> A polypeptide encoded by SEQ ID NO:9

<400> 20
 Met Asn Tyr Lys Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Thr Gln
 1 5 10 15
 20 Lys His Leu Ile His Gly Gly Glu Gly Leu Phe Gln His Glu Leu Arg
 20 25 30
 Ala Val Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
 35 40 45
 Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
 25 50 55 60
 Ile Val Ser Arg Gln Ser Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
 65 70 75 80
 Cys Arg His Arg Gly Lys Thr Leu Val Asn Ala Glu Ala Gly Asn Ala
 85 90 95
 30 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
 100 105 110
 Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ser Leu
 115 120 125
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 35 130 135 140
 His Gly Phe Ile Tyr Ala Cys Ile Asp Gln Glu Ala Pro Ser Leu Met
 145 150 155 160
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His
 165 170 175
 40 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys

| | | | |
|----|---|-----|-----|
| | 180 | 185 | 190 |
| | Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His | | |
| | 195 | 200 | 205 |
| | Val Gly Trp Thr His Ala Ser Ser Leu Cys Thr Gly Glu Ser Ile Phe | | |
| 5 | 210 | 215 | 220 |
| | Ser Ser Leu Ala Gly Asn Ala Val Leu Pro Pro Glu Gly Ala Gly Leu | | |
| | 225 | 230 | 235 |
| | Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly | | |
| | 245 | 250 | 255 |
| 10 | Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe | | |
| | 260 | 265 | 270 |
| | Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg | | |
| | 275 | 280 | 285 |
| | Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn | | |
| 15 | 290 | 295 | 300 |
| | Ser Val Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp | | |
| | 305 | 310 | 315 |
| | Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met | | |
| | 325 | 330 | 335 |
| 20 | Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val | | |
| | 340 | 345 | 350 |
| | Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr | | |
| | 355 | 360 | 365 |
| | Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Ile | | |
| 25 | 370 | 375 | 380 |
| | Ser Asn Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Ala Val Tyr Pro | | |
| | 385 | 390 | 395 |
| | Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe | | |
| | 405 | 410 | 415 |
| 30 | Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe | | |
| | 420 | 425 | 430 |
| | Glu Asp Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp | | |
| | 435 | 440 | 445 |

Arg

35

<210> 21

<211> 449

<212> PRT

40

<213> Artificial Sequence

<220>

<223> A polypeptide encoded by SEQ ID NO:10

<400> 21

```

5 Met Asn Tyr Lys Asn Lys Asn Leu Val Ser Glu Ser Gly Leu Thr Gln
  1             5             10             15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln Arg Glu Leu Glu
      20             25             30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
10      35             40             45
Pro Ser Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val
      50             55             60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65             70             75             80
15 Cys Arg His Arg Gly Lys Thr Leu Val His Ala Glu Ala Gly Asn Ala
      85             90             95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ala Asn Gly
      100            105            110
Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Glu Ala Leu
20      115            120            125
Asp Lys Lys Cys Met Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
      130            135            140
His Gly Phe Ile Tyr Gly Cys Phe Asp Glu Glu Ala Pro Ser Leu Lys
      145            150            155            160
25 Asp Tyr Met Gly Asp Ala Gly Trp Tyr Leu Glu Pro Met Phe Lys His
      165            170            175
Ser Gly Gly Leu Glu Leu Ile Gly Pro Pro Gly Lys Val Ile Ile Lys
      180            185            190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Thr Gly Asp Ala Tyr His
30      195            200            205
Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Gln Ser Val Phe
      210            215            220
Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
225            230            235            240
35 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
      245            250            255
Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
      260            265            270
Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Glu Val Arg
40      275            280            285

```

Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 Ser Phe Leu Thr Cys Ser Gly Val Phe Lys Val Trp His Pro Ile Asp
 305 310 315 320
 5 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Met Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Val Asp Ala Val Gln Arg Thr Val
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 10 355 360 365
 Val Ser Gln Asn Ala Lys Lys Tyr Gln Ser Arg Asp Gly Asp Leu Val
 370 375 380
 Ser Asn Leu Gly Phe Gly Gly Asp Val Tyr Gly Asp Glu Val Tyr Pro
 385 390 395 400
 15 Gly Ile Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Gly Ala His Ile Ser Ser Ser Ser Trp Ala Glu Phe
 420 425 430
 Glu Asp Val Ser Lys Asn Trp His Thr Glu Leu Ala Lys Thr Thr Asp
 20 435 440 445
 Arg

<210> 22
 25 <211> 447
 <212> PRT
 <213> Artificial Sequence
 <220>
 30 <223> A polypeptide encoded by SEQ ID NO:11

<400> 22
 Met Ile Tyr Glu Asn Leu Val Ser Glu Ala Gly Leu Thr Gln Lys His
 1 5 10 15
 35 Leu Ile His Gly Asp Lys Glu Leu Phe Gln His Glu Leu Lys Thr Ile
 20 25 30
 Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile Pro Ser
 35 40 45
 Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Val Asp Glu Val Ile Val
 40 50 55 60

Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe Leu Asn Val Cys Arg
 65 70 75 80
 His Arg Gly Lys Thr Leu Val His Ala Glu Ala Gly Asn Ala Lys Gly
 85 90 95
 5 Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly Glu Leu
 100 105 110
 Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly Asp Thr Ile Lys Lys
 115 120 125
 Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile Glu Ser Phe His Gly
 10 130 135 140
 Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro Thr Leu Val Asp Tyr
 145 150 155 160
 Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Ile Phe Lys His Ser Gly
 165 170 175
 15 Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys Ala Asn
 180 185 190
 Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His Val Gly
 195 200 205
 Trp Thr His Ala Ser Ser Leu Arg Ser Gly Gln Ser Ile Phe Thr Pro
 20 210 215 220
 Leu Ala Gly Asn Ala Met Leu Pro Pro Glu Gly Ala Gly Leu Gln Met
 225 230 235 240
 Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly Tyr Ser
 245 250 255
 25 Gly Val His Ser Ala Asp Leu Val Pro Glu Met Met Ala Phe Gly Gly
 260 265 270
 Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly Asp Val Arg Ala Arg
 275 280 285
 Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn Ser Ile
 30 290 295 300
 Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp Glu Asn
 305 310 315 320
 Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met Pro Glu
 325 330 335
 35 Asp Leu Lys Arg Arg Leu Ala Asp Ala Val Gln Arg Thr Val Gly Pro
 340 345 350
 Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr Glu Ser
 355 360 365
 Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp Leu Ile Ala Asn
 40 370 375 380

| | | | | | | | | | | | | | | | | |
|-----|-----|-------|---------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Gly | Phe | Gly | Lys | Asp | Val | Tyr | Gly | Asp | Glu | Cys | Tyr | Pro | Gly | Val | |
| 385 | | | | | | 390 | | | | | 395 | | | | 400 | |
| Val | Ala | Lys | Ser | Ala | Ile | Gly | Glu | Thr | Ser | Tyr | Arg | Gly | Phe | Tyr | Arg | |
| | | | | 405 | | | | | 410 | | | | | 415 | | |
| 5 | Ala | Tyr | Gln | Ala | His | Ile | Ser | Ser | Ser | Asn | Trp | Ala | Glu | Phe | Glu | Asn |
| | | | | 420 | | | | | 425 | | | | | 430 | | |
| | Thr | Ser | Arg | Asn | Trp | His | Thr | Glu | Leu | Thr | Lys | Thr | Thr | Asp | Arg | |
| | | | 435 | | | | | 440 | | | | | 445 | | | |
| 10 | | <210> | 23 | | | | | | | | | | | | | |
| | | <211> | 447 | | | | | | | | | | | | | |
| | | <212> | PRT | | | | | | | | | | | | | |
| | | <213> | Artificial Sequence | | | | | | | | | | | | | |
| 15 | | <220> | | | | | | | | | | | | | | |
| | | <223> | A polypeptide encoded by SEQ ID NO:12 | | | | | | | | | | | | | |
| | | <400> | 23 | | | | | | | | | | | | | |
| Met | Ser | Tyr | Gln | Asn | Leu | Val | Ser | Glu | Ala | Gly | Leu | Thr | Gln | Lys | Leu | |
| 20 | 1 | | | 5 | | | | 10 | | | | | | 15 | | |
| Leu | Ile | His | Gly | Asp | Lys | Glu | Leu | Phe | Gln | His | Glu | Leu | Lys | Thr | Ile | |
| | | | 20 | | | | | 25 | | | | | 30 | | | |
| Phe | Ala | Arg | Asn | Trp | Leu | Phe | Leu | Thr | His | Asp | Ser | Leu | Ile | Pro | Ser | |
| | | 35 | | | | 40 | | | | | | 45 | | | | |
| 25 | Pro | Gly | Asp | Tyr | Val | Thr | Ala | Lys | Met | Gly | Val | Asp | Glu | Val | Ile | Val |
| | | 50 | | | | 55 | | | | | | 60 | | | | |
| Ser | Arg | Gln | Asn | Asp | Gly | Ser | Val | Arg | Ala | Phe | Leu | Asn | Val | Cys | Arg | |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | | |
| His | Arg | Gly | Lys | Thr | Leu | Val | His | Thr | Glu | Ala | Gly | Asn | Ala | Lys | Gly | |
| 30 | | | | 85 | | | | | 90 | | | | | 95 | | |
| Phe | Val | Cys | Gly | Tyr | His | Gly | Trp | Gly | Tyr | Gly | Ser | Asn | Gly | Glu | Leu | |
| | | | 100 | | | | | 105 | | | | | 110 | | | |
| Gln | Ser | Val | Pro | Phe | Glu | Lys | Glu | Leu | Tyr | Gly | Asp | Ala | Ile | Lys | Lys | |
| | | | 115 | | | | | 120 | | | | | 125 | | | |
| 35 | Lys | Cys | Leu | Gly | Leu | Lys | Glu | Val | Pro | Arg | Ile | Glu | Ser | Phe | His | Gly |
| | | 130 | | | | | | 135 | | | | | 140 | | | |
| Phe | Ile | Tyr | Gly | Cys | Phe | Asp | Ala | Glu | Ala | Pro | Pro | Leu | Ile | Asp | Tyr | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | |
| Leu | Gly | Asp | Ala | Ala | Trp | Tyr | Leu | Glu | Pro | Thr | Phe | Lys | His | Ser | Gly | |
| 40 | | | | 165 | | | | | 170 | | | | | | 175 | |

Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Val Lys Ala Asn
 180 185 190
 Trp Lys Pro Phe Ala Glu Asn Phe Val Gly Asp Ile Tyr His Val Gly
 195 200 205
 5 Trp Thr His Ala Ala Ala Leu Arg Ala Gly Gln Ser Val Phe Ser Ser
 210 215 220
 Leu Ala Gly Asn Ala Lys Leu Pro Pro Glu Gly Ala Gly Leu Gln Met
 225 230 235 240
 Thr Ser Lys Tyr Gly Ser Gly Met Gly Leu Thr Trp Asp Tyr Tyr Ser
 10 245 250 255
 Gly Asn Phe Ser Ala Asp Met Val Pro Asp Leu Met Ala Phe Gly Ala
 260 265 270
 Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly Asp Val Arg Ala Arg
 275 280 285
 15 Ile Tyr Arg Ser Ile Leu Asn Gly Thr Val Phe Pro Asn Asn Ser Phe
 290 295 300
 Leu Thr Gly Ser Ala Thr Phe Lys Val Trp Asn Pro Ile Asp Glu Asn
 305 310 315 320
 Thr Thr Glu Val Trp Thr Tyr Ala Phe Val Glu Lys Asp Met Pro Glu
 20 325 330 335
 Asp Leu Lys Arg Arg Leu Ala Asp Ala Ala Gln Arg Ser Val Gly Pro
 340 345 350
 Ala Gly Phe Trp Glu Ser Asp Asp Asn Glu Asn Met Glu Thr Leu Ser
 355 360 365
 25 Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp Gln Ile Ala Ser
 370 375 380
 Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys Tyr Pro Gly Val
 385 390 395 400
 Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe Tyr Arg
 30 405 410 415
 Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala Glu Phe Glu Asn
 420 425 430
 Ala Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr Thr Asp Arg
 435 440 445
 35
 <210> 24
 <211> 451
 <212> PRT
 <213> Artificial Sequence

<220>

<223> A polypeptide encoded by SEQ ID NO:13

<400> 24

```

5 Met Arg Gln Ala Ile Met Ser Tyr Gln Asn Leu Val Ser Glu Ala Gly
  1             5             10             15
  Leu Thr Gln Lys His Leu Ile Tyr Gly Asp Lys Glu Leu Phe Gln His
    20             25             30
  Glu Leu Lys Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp
10      35             40             45
  Ser Leu Ile Pro Ser Pro Gly Asp Tyr Val Lys Ala Lys Met Gly Val
    50             55             60
  Asp Glu Val Ile Val Ser Arg Gln Asn Asp Gly Ser Val Arg Ala Phe
  65             70             75             80
15 Leu Asn Val Cys Arg His Arg Gly Lys Thr Ile Val Asp Ala Glu Ala
    85             90             95
  Gly Asn Ala Lys Gly Phe Val Cys Gly Tyr His Gly Trp Gly Tyr Gly
    100            105            110
  Ser Asn Gly Glu Leu Gln Ser Val Pro Phe Glu Lys Glu Leu Tyr Gly
20      115            120            125
  Asp Ala Ile Lys Lys Lys Cys Leu Gly Leu Lys Glu Val Pro Arg Ile
    130            135            140
  Glu Ser Phe His Gly Phe Ile Tyr Gly Cys Phe Asp Ala Glu Ala Pro
  145            150            155            160
25 Pro Leu Ile Asp Tyr Leu Gly Asp Val Ala Trp Tyr Leu Glu Pro Thr
    165            170            175
  Phe Lys His Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Ala Lys Val
    180            185            190
  Val Val Lys Gly Asn Trp Lys Val Phe Ala Glu Asn Phe Val Gly Asp
30      195            200            205
  Ile Tyr His Ile Gly Trp Thr His Ala Ser Ile Leu Arg Ala Gly Gln
    210            215            220
  Ala Ile Phe Ala Pro Leu Ala Gly Asn Ala Met Leu Pro Pro Glu Gly
  225            230            235            240
35 Thr Gly Leu Gln Ala Thr Thr Lys Tyr Gly Ser Gly Ile Gly Val Ser
    245            250            255
  Leu Asp Ala Tyr Ser Gly Val Gln Ser Ala Asp Leu Val Pro Glu Met
    260            265            270
  Met Ala Phe Gly Gly Ala Lys Gln Glu Lys Leu Ala Lys Glu Ile Gly
40      275            280            285

```

Asp Val Arg Ala Arg Ile Tyr Arg Ser Gln Val Asn Gly Thr Val Phe
 290 295 300
 Pro Asn Asn Cys Phe Leu Thr Gly Ala Gly Val Phe Lys Val Phe Asn
 305 310 315 320
 5 Pro Ile Asp Glu Asn Thr Thr Glu Ala Trp Thr Tyr Ala Ile Val Glu
 325 330 335
 Lys Asp Met Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ala Ala Gln
 340 345 350
 Arg Ser Val Gly Pro Ala Gly Tyr Trp Glu Ser Asp Asp Asn Asp Asn
 10 355 360 365
 Met Val Leu Ser Gln Asn Ala Lys Lys Tyr Gln Ser Ser Asn Ser Asp
 370 375 380
 Leu Ile Ala Asp Leu Gly Phe Gly Lys Asp Val Tyr Gly Asp Glu Cys
 385 390 395 400
 15 Tyr Pro Gly Val Val Ser Lys Ser Ala Phe Ser Glu Thr Asn His Arg
 405 410 415
 Gly Phe Tyr Arg Ala Tyr Gln Ala His Ile Ser Ser Ser Asn Trp Ala
 420 425 430
 Glu Phe Glu Asn Thr Ser Arg Asn Trp His Thr Glu Leu Thr Lys Thr
 20 435 440 445
 Thr Asp Arg
 450

<210> 25

25 <211> 2265

<212> DNA

<213> Pseudomonas sp.

<400> 25

30 gagggtagag aaatcgaatg ccccttgcat caaggctcggg ttgacgtttg cacaggcaaa 60
 gccctgtgcg caccctgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120
 cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt 180
 aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240
 aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca 300
 35 tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360
 tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga 420
 gaacgacggg tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggg 480
 gagcgtggaa gccggcaatg ccaaagggtt tgtttgacgc tatcacggct ggggcttcgg 540
 ctccaacggg gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa 600
 40 taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta 660

cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtta 720
 cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt 780
 tgtgatcaag gccaaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt 840
 gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg 900
 5 caatgcggcg ctaccacctg aaggcgcagg cttgcaaagt acctccaaat acggcagcgg 960
 catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020
 gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc 1080
 tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
 ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacctg 1200
 10 cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca 1260
 gcgaacgttc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
 ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380
 cggtgaggac gtatacggcg acgcggtcta tccaggcgct gtcggcaaat cggcgatcgg 1440
 cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg 1500
 15 ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560
 ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620
 gacgcgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
 actacgctgc tgaccagga agcgcatctg ttggacattc aggettaccg tgcttggtta 1740
 gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
 20 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
 ctgaaagtcc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
 cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
 cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100
 25 cagcgattcg tcgattacc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
 tgattcagtg accattttta caaatggta ctgcaaccgc ggtcaccatt aatcaaaggg 2220
 aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg 2265

<210> 26

30 <211> 449

<212> PRT

<213> Pseudomonas sp.

<400> 26

35 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln

1 5 10 15

Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys

20 25 30

Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile

40 35 40 45

Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
 50 55 60
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
 65 70 75 80
 5 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
 85 90 95
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
 100 105 110
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
 10 115 120 125
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 130 135 140
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
 145 150 155 160
 15 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
 165 170 175
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
 180 185 190
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
 20 195 200 205
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
 210 215 220
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
 225 230 235 240
 25 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
 245 250 255
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 30 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 35 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Phe
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 40 355 360 365

Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 370 375 380
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 5 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 10 435 440 445
 Arg

<210> 27

15 <211> 2265

<212> DNA

<213> Artificial Sequence

<220>

20 <223> A modified DNA sequence

<400> 27

| | | |
|----|--|------|
| | gagggtagag aaatcgaatg ccccttgcat caaggctcggg ttgacgtttg cacaggcaaa | 60 |
| | gccctgtgcg caccgtgac acagaacatc aaaacatcag cagtcaagat tgagaacctg | 120 |
| 25 | cgcgtaaatga ttgatttgag ctaagaattt taacaggagg caccgccggc cctagagcgt | 180 |
| | aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt | 240 |
| | aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca | 300 |
| | tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcttgattcc | 360 |
| | tgcccccgcc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgcca | 420 |
| 30 | gaacgaecgt tcgattcgtg cttttctgaa cgtttgccgg catcgtggca agacgctggg | 480 |
| | gagcgtggaa gccggcaatg ccaaagggtt tgtttgacgc tatcacggct ggggcttcgg | 540 |
| | ctccaacggg gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa | 600 |
| | taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta | 660 |
| | cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg | 720 |
| 35 | cctggaacct atgttcaagc attccggcgg tttagaactg gtcggctctc caggcaagg | 780 |
| | tgtgatcaag gccaaactga aggcacccgc ggaaaaactt gtgggagatg cataccacgt | 840 |
| | gggttggaag cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg | 900 |
| | caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg | 960 |
| | catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt | 1020 |
| 40 | gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc | 1080 |

tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
 ctccgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggaccta 1200
 cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttggccg actctgttca 1260
 gcgaacgggc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
 5 ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380
 cgggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtccggcaaat cggcgatcgg 1440
 cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cagtcagca gctccaactg 1500
 ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560
 ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620
 10 gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
 actacgctgc tgaccagga agcgcatctt ttggacattc aggttaccg tgcttggtta 1740
 gagcactgcg tggggtcaga ggtgcaatat caggtcattt cagcggaact gcgcgcagct 1800
 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
 ctgaaagtgc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
 15 cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaata caaagagcta 1980
 cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100
 cagcgattcg tcgattacc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
 tgattcagtg accattttta caaatggtca ctgcaaccgc gggtcaccatt aatcaaaggg 2220
 20 aatgtacgtg tatgggcaat caacaagtcg ttctcgataac cgggtg 2265

<210> 28

<211> 2265

<212> DNA

25 <213> Artificial Sequence

<220>

<223> A modified DNA sequence

30 <400> 28

gagggtagag aaatcgaatg ccccttgcac caaggtcggt ttgacgtttg cacaggcaaa 60
 gccctgtgcg caccctgcac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120
 cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt 180
 aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggc 240
 35 aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca 300
 tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360
 tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga 420
 gaacgacggc tcgattcgtg cttttctgaa cgtttgccg catcgtggca agacgctggt 480
 gagcgtggaa gccggcaatg ccaaagggtt tgtttgacgc tatcacggct ggggcttcgg 540
 40 ctccaacggc gaactgcaga gcgttcatt tgaaaaagat ctgtacggcg agtcgctcaa 600

```

taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcaccta 660
cgggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg 720
cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt 780
tgtgatcaag gccaaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt 840
5 gggttggacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg 900
caatgcggcg ctaccacctg aaggcgcagg cttgcaaata acctccaaat acggcagcgg 960
catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020
gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc 1080
tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
10 ctcggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacctg 1200
cgccattgtc gaaaaagaca tgccctgagga tctcaagcgc cgcttgggcg actctgttca 1260
gcgaacggcc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttgggtt 1380
cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440
15 cgagaccagt tatcgtggtt tctaccgggc ttaccaggca cacgtcagca gctccaactg 1500
ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg 1560
ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
actacgctgc tgaccaggga agcgcatctg ttggacattc aggccttaccg tgcttggtta 1740
20 gagcactcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
ctgaaagtgc gagttgagca tcaactggat ccgcaaaaact ggggcaacag cccgaagctg 1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc 2040
25 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100
cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg 2265

```

```

30 <210> 29
    <211> 2265
    <212> DNA
    <213> Artificial Sequence

```

```

35 <220>
    <223> A modified DNA sequence

```

```

    <400> 29
gagggtagag aaatcgaatg ccccttgcac caaggtcggt ttgacgtttg cacaggcaaa 60
40 gccctgtgcg caccctgtgac acagaacatc aaaacatatc cagtcaagat tgagaacctg 120

```

| | | |
|----|--|------|
| | cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccgccggc cctagagcgt | 180 |
| | aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt | 240 |
| | aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca | 300 |
| | tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc | 360 |
| 5 | tgcccccggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga | 420 |
| | gaacgacggg tcgattcggt cttttctgaa cgtttgccgg catcgaggca agacgctggg | 480 |
| | gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg | 540 |
| | ctccaacggg gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa | 600 |
| | taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta | 660 |
| 10 | cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggt | 720 |
| | cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaagg | 780 |
| | tgtgatcaag gccaaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt | 840 |
| | gggttgagcg cagcgtctt cgcttcgctc gggggagctc atcttctcgt cgctcgctgg | 900 |
| | caatgcggcg ctaccacctg aaggcgcagg cttgcaaag acctccaaat acggcagcgg | 960 |
| 15 | catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt | 1020 |
| | gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc | 1080 |
| | tccgatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgetgacctg | 1140 |
| | ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacct | 1200 |
| | cgccattgtc gaaaaagaca tgcctgagga tctcaagcgc cgcttgcccg actctgttca | 1260 |
| 20 | gcgaacgacc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc | 1320 |
| | ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggtt | 1380 |
| | cggtagggac gtatacggcg acgcgggtcta tccaggcgtc gtcggcaaat cggcgatcgg | 1440 |
| | cgagaccagt tatcgtgggt tctaccgggc ttaccaggca cagtcagca gctccaactg | 1500 |
| | ggctgagttc gagcatgcct ctactacttg gcatactgaa cttacgaaga ctactgatcg | 1560 |
| 25 | ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac | 1620 |
| | gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc | 1680 |
| | actacgctgc tgaccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta | 1740 |
| | gagcactgcg tggggtcaga ggtgcaatat caggtcattt cagcgaaact gcgcgagct | 1800 |
| | tcagagcgct gtataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagca | 1860 |
| 30 | ctgaaagtgc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg | 1920 |
| | cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta | 1980 |
| | cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc | 2040 |
| | ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc | 2100 |
| | cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg | 2160 |
| 35 | tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg | 2220 |
| | aatgtacgtg tatgggcaat caacaagtcg tttcgataac cgggtg | 2265 |

<210> 30

<211> 2265

40

<212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

5

<400> 30

| | | |
|----|--|------|
| | gagggtagag aaatcgaatg ccccttgcac caaggctcgtt ttgacgtttg cacaggcaaa | 60 |
| | gccctgtgcg caccctgtac acagaacatc aaaacatata cagtcaagat tgagaacctg | 120 |
| | cgcgtaatga ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt | 180 |
| 10 | aatcaccctt attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggg | 240 |
| | aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca | 300 |
| | tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc | 360 |
| | tgcccccgcc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctcccgga | 420 |
| | gaacgacggg tcgattcgtg cttttctgaa cgtttgccg catcgtggca agacgctggg | 480 |
| 15 | gagcgtggaa gccggcaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg | 540 |
| | ctccaacggg gaactgcaga gcgttccatt tgaaaaagat ctgtacggcg agtcgctcaa | 600 |
| | taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gcttcatcta | 660 |
| | cggttgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggt | 720 |
| | cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaagg | 780 |
| 20 | tgtgatcaag gccaactgga aggcacccgc ggaactttt gtgggagatg cataccacgt | 840 |
| | gggttgagcg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg | 900 |
| | caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg | 960 |
| | catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccgggaatt | 1020 |
| | gatggcattc ggaggcgcaa agcaggaaag gctgaacaaa gaaattggcg atgttcgcgc | 1080 |
| 25 | tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg | 1140 |
| | ctcgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacct | 1200 |
| | cgcattgtc gaaaaagaca tgcttgagga tctcaagcgc cgcttggccg actctgttca | 1260 |
| | gcgaacgctc gggcctgctg gcttctggga aagcgcgcac aatgacaata tggaaacagc | 1320 |
| | ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt | 1380 |
| 30 | cggtgaggac gtatacggcg acgcgttcta tccaggcgtc gtcggcaaat cggcgatcgg | 1440 |
| | cgagaccagt tatcgtgggt tctaccgggc ttaccaggca cacgtcagca gctccaactg | 1500 |
| | ggctgagttc gagcatgcct ctagtacttg gcatactgaa cttacgaaga ctactgatcg | 1560 |
| | ctaacagacg agtcgacct gatgatcaat attcaagaag acaagctggg ttccgcccac | 1620 |
| | gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc | 1680 |
| 35 | actacgtgc tgaccagga agcgcatttg ttggacattc aggcttaccg tgcttggtta | 1740 |
| | gagcactgcg tggggtcaga ggtgcaatat caggctcatt cacgcgaact gcgcgcagct | 1800 |
| | tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa | 1860 |
| | ctgaaagtgc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg | 1920 |
| | cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta | 1980 |
| 40 | cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgtc | 2040 |

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| ttctacgccg | cccggaaga | taaattgaaa | cgtggcgaag | gtggagtacg | aaaattggtc | 2100 |
| cagcgattcg | tcgattaccc | agagcgcata | cttcagacgc | acaatctgat | ggtctttctg | 2160 |
| tgattcagtg | accattttta | caaattggta | ctgcaaccgc | ggtcaccatt | aatcaaaggg | 2220 |
| aatgtacgtg | tatgggcaat | caacaagtcg | tttcgataac | cggtg | | 2265 |

5

<210> 31

<211> 2265

<212> DNA

<213> Artificial Sequence

10

<220>

<223> A modified DNA sequence

<400> 31

| | | | | | | | |
|----|-------------|------------|-------------|------------|------------|-------------|------|
| 15 | gagggtagag | aaatcgaatg | ccccttgcac | caaggtcggt | ttgacgtttg | cacaggcaaa | 60 |
| | gccctgtgcg | cacccgtgac | acagaacatc | aaaacatata | cagtcaagat | tgagaacctg | 120 |
| | cgcgtaatga | ttgatttgag | ctaagaattt | taacaggagg | caccccgggc | cctagagcgt | 180 |
| | aatcaccccc | attccatctt | ttttaggtga | aaacatgaat | tacaataata | aaatcttggg | 240 |
| | aagtgaatct | ggtctgagcc | aaaagcacct | gattcatggc | gatgaagaac | ttttccaaca | 300 |
| 20 | tgaactgaaa | accatttttg | cgcggaactg | gctttttctc | actcatgata | gcctgattcc | 360 |
| | tgcccccggc | gactatgtta | ccgcaaaaat | ggggattgac | gaggtcatcg | tctcccggca | 420 |
| | gaacgacggt | tcgattcgtg | cttttctgaa | cgtttgccgg | catcgtggca | agacgctggg | 480 |
| | gagcgtggaa | gccggcaatg | ccaaagggtt | tgtttgacgc | tatcacggct | ggggcttcgg | 540 |
| | ctccaacggt | gaactgcaga | gcgttccatt | tgaaaaagat | ctgtacggcg | agtcgctcaa | 600 |
| 25 | taaaaaatgt | ctgggggtga | aagaagtcgc | tcgcgtggag | agcttccatg | gcttcatcta | 660 |
| | cggttgcttc | gaccaggagg | cccctcctct | tatggactat | ctgggtgacg | ctgcttggtg | 720 |
| | cctggaacct | atgttcaagc | attccggcgg | tttagaactg | gtcggtcctc | caggcaagggt | 780 |
| | tgtgatcaag | gccaactgga | aggcaccgcg | ggaaaacttt | gtgggagatg | cataccacgt | 840 |
| | gggttggaag | cacgcgtctt | cgcttcgctc | gggggagtc | atcttctcgt | cgctcgctgg | 900 |
| 30 | caatgcggcg | ctaccacctg | aaggcgcagg | cttgcaaatg | acctccaaat | acggcagcgg | 960 |
| | catgggtgtg | ttgtgggacg | gatattcagg | tgtgcatagc | gcagacttgg | ttccggaatt | 1020 |
| | gatggcattc | ggaggcgcaa | agcaggaaag | gctgaacaaa | gaaattggcg | atgttcgcgc | 1080 |
| | tcggatttat | cgcagccacc | tcaactgcac | cgttttcccg | aacaacagca | tgctgacctg | 1140 |
| | ctcgggtgtt | ttcaaagtat | ggaaccgat | cgacgcaaac | accaccgagg | tctggacctg | 1200 |
| 35 | cgccattgtc | gaaaaagaca | tgccctgagga | tctcaagcgc | cgcttggccg | actctgttca | 1260 |
| | gcgaacgatc | gggcctgctg | gcttctggga | aagcgacgac | aatgacaata | tggaacagc | 1320 |
| | ttcgcaaaac | ggcaagaaat | atcaatcaag | agatagtgat | ctgctttcaa | accttggttt | 1380 |
| | cggtagaggac | gtatacggcg | acgcggtcta | tccaggcgct | gtcggcaaat | cggcgatcgg | 1440 |
| | cgagaccagt | tatcgtgggt | tctaccgggc | ttaccaggca | cacgtcagca | gctccaactg | 1500 |
| 40 | ggctgagttc | gagcatgcct | ctagtacttg | gcatactgaa | cttacgaaga | ctactgatcg | 1560 |

```

ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggg ttccgcccac 1620
gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
actacgctgc tgaccagga agcgcatctg ttggacattc aggcttaccg tgcttggtta 1740
gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
5 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
ctgaaagtgc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgct 2040
ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggct 2100
10 cagcgattcg tcgattacc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220
aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg 2265

```

<210> 32

15 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

20 <223> A polypeptide encoded by SEQ ID NO:27

<400> 32

```

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1             5             10             15
25 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
    20             25             30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
    35             40             45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
30 50             55             60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65             70             75             80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
    85             90             95
35 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
    100             105             110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
    115             120             125
Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
40 130             135             140

```

His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
 145 150 155 160
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
 165 170 175
 5 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
 180 185 190
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
 195 200 205
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
 10 210 215 220
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
 225 230 235 240
 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
 245 250 255
 15 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 20 290 295 300
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 25 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Gly
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 355 360 365
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 30 370 375 380
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 35 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 435 440 445

Arg

<210> 33
 <211> 449
 <212> PRT
 <213> Artificial Sequence

5

<220>
 <223> A polypeptide encoded by SEQ ID NO:28

<400> 33

```

10 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
    1             5             10             15
    Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
        20             25             30
    Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
15      35             40             45
    Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
        50             55             60
    Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
    65             70             75             80
20 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
        85             90             95
    Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
        100            105            110
    Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
25      115            120            125
    Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
        130            135            140
    His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
    145            150            155            160
30 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
        165            170            175
    Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
        180            185            190
    Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
35      195            200            205
    Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
        210            215            220
    Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
    225            230            235            240
40 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
  
```


245 250 255
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 5 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 10 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Ala
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 15 355 360 365
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 370 375 380
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 20 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 25 435 440 445
 Arg

<210> 34

30 <211> 449

<212> PRT

<213> Artificial Sequence

<220>

35 <223> A polypeptide encoded by SEQ ID NO:29

<400> 34

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1 5 10 15
 40 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys

| | | | | | |
|----|---|-----|-----|-----|-----|
| | 20 | | 25 | | 30 |
| | Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile | | | | |
| | 35 | | 40 | | 45 |
| | Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val | | | | |
| 5 | 50 | | 55 | | 60 |
| | Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val | | | | |
| | 65 | | 70 | | 75 |
| | Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala | | | | |
| | | 85 | | 90 | 95 |
| 10 | Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly | | | | |
| | 100 | | 105 | | 110 |
| | Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu | | | | |
| | 115 | | 120 | | 125 |
| | Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe | | | | |
| 15 | 130 | | 135 | | 140 |
| | His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met | | | | |
| | 145 | | 150 | | 155 |
| | Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His | | | | |
| | | 165 | | 170 | 175 |
| 20 | Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys | | | | |
| | 180 | | 185 | | 190 |
| | Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His | | | | |
| | 195 | | 200 | | 205 |
| | Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe | | | | |
| 25 | 210 | | 215 | | 220 |
| | Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu | | | | |
| | 225 | | 230 | | 235 |
| | Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly | | | | |
| | | 245 | | 250 | 255 |
| 30 | Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe | | | | |
| | 260 | | 265 | | 270 |
| | Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg | | | | |
| | 275 | | 280 | | 285 |
| | Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn | | | | |
| 35 | 290 | | 295 | | 300 |
| | Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp | | | | |
| | 305 | | 310 | | 315 |
| | Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met | | | | |
| | | 325 | | 330 | 335 |
| 40 | Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Thr | | | | |

```

      340      345      350
Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
      355      360      365
Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
5   370      375      380
Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
385      390      395      400
Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
      405      410      415
10 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
      420      425      430
Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
      435      440      445
Arg
15
      <210> 35
      <211> 449
      <212> PRT
20   <213> Artificial Sequence

      <220>
      <223> A polypeptide encoded by SEQ ID NO:30

25   <400> 35
Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
  1           5           10           15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
      20           25           30
30 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
      35           40           45
Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
      50           55           60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
35 65           70           75           80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
      85           90           95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
      100          105          110
40 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu

```

| | | | | | |
|----|---|-----|-----|-----|-----|
| | 115 | | 120 | | 125 |
| | Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe | | | | |
| | 130 | | 135 | | 140 |
| | His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met | | | | |
| 5 | 145 | | 150 | | 155 |
| | Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His | | | | |
| | | 165 | | 170 | 175 |
| | Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys | | | | |
| | 180 | | 185 | | 190 |
| 10 | Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His | | | | |
| | 195 | | 200 | | 205 |
| | Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe | | | | |
| | 210 | | 215 | | 220 |
| | Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu | | | | |
| 15 | 225 | | 230 | | 235 |
| | Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly | | | | |
| | | 245 | | 250 | 255 |
| | Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe | | | | |
| | 260 | | 265 | | 270 |
| 20 | Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg | | | | |
| | 275 | | 280 | | 285 |
| | Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn | | | | |
| | 290 | | 295 | | 300 |
| | Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp | | | | |
| 25 | 305 | | 310 | | 315 |
| | Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met | | | | |
| | | 325 | | 330 | 335 |
| | Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Leu | | | | |
| | 340 | | 345 | | 350 |
| 30 | Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr | | | | |
| | 355 | | 360 | | 365 |
| | Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu | | | | |
| | 370 | | 375 | | 380 |
| | Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro | | | | |
| 35 | 385 | | 390 | | 395 |
| | Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe | | | | |
| | | 405 | | 410 | 415 |
| | Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe | | | | |
| | 420 | | 425 | | 430 |
| 40 | Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp | | | | |

435

440

445

Arg

5 <210> 36

<211> 449

<212> PRT

<213> Artificial Sequence

10 <220>

<223> A polypeptide encoded by SEQ ID NO:31

<400> 36

```

Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
15 1           5           10           15
Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
           20           25           30
Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
           35           40           45
20 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
           50           55           60
Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
65           70           75           80
Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
25           85           90           95
Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
           100          105          110
Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
           115          120          125
30 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
           130          135          140
His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
145          150          155          160
Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
35          165          170          175
Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
           180          185          190
Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
           195          200          205
40 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe

```

| | | | |
|----|---|-----|---------|
| | 210 | 215 | 220 |
| | Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu | | |
| | 225 | 230 | 235 240 |
| | Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly | | |
| 5 | 245 | 250 | 255 |
| | Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe | | |
| | 260 | 265 | 270 |
| | Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg | | |
| | 275 | 280 | 285 |
| 10 | Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn | | |
| | 290 | 295 | 300 |
| | Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp | | |
| | 305 | 310 | 315 320 |
| | Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met | | |
| 15 | 325 | 330 | 335 |
| | Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Ile | | |
| | 340 | 345 | 350 |
| | Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr | | |
| | 355 | 360 | 365 |
| 20 | Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu | | |
| | 370 | 375 | 380 |
| | Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro | | |
| | 385 | 390 | 395 400 |
| | Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe | | |
| 25 | 405 | 410 | 415 |
| | Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe | | |
| | 420 | 425 | 430 |
| | Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp | | |
| | 435 | 440 | 445 |

30 Arg

<210> 37

<211> 22

35 <212> DNA

<213> Artificial Sequence

<220>

<223> An oligonucleotide

40

| | | |
|----|--------------------------------------|----|
| | <400> 37 | |
| | ttcagcgaac ggtcgggcct gc | 22 |
| | <210> 38 | |
| 5 | <211> 27 | |
| | <212> DNA | |
| | <213> Artificial Sequence | |
| | <220> | |
| 10 | <223> An oligonucleotide | |
| | <400> 38 | |
| | gccgggcctc ttgcgggata tcgtcca | 27 |
| 15 | <210> 39 | |
| | <211> 27 | |
| | <212> DNA | |
| | <213> Artificial Sequence | |
| 20 | <220> | |
| | <223> An oligonucleotide | |
| | <400> 39 | |
| | gttgccattg ctgcaggcat cgtggtg | 27 |
| 25 | <210> 40 | |
| | <211> 33 | |
| | <212> DNA | |
| | <213> Artificial Sequence | |
| 30 | <220> | |
| | <223> An oligonucleotide | |
| | <400> 40 | |
| 35 | gaggcacccg cggaagcttt tgtgggagat gca | 33 |
| | <210> 41 | |
| | <211> 30 | |
| | <212> DNA | |
| 40 | <213> Artificial Sequence | |

<220>

<223> An oligonucleotide

<400> 41

5 gcacccgcgg aacaatttgt gggagatgca

30

<210> 42

<211> 21

<212> DNA

10 <213> Artificial Sequence

<220>

<223> An oligonucleotide

15 <400> 42

ccgcggaaaag ctttgtggga g

21

<210> 43

<211> 24

20 <212> DNA

<213> Artificial Sequence

<220>

<223> An oligonucleotide

25

<400> 43

ccgcggaaaa gcttgtggga gatg

24

<210> 44

30 <211> 23

<212> DNA

<213> Artificial Sequence

<220>

35 <223> An oligonucleotide

<400> 44

cgcggaac gttgtgggag atg

23

40 <210> 45

<211> 23
 <212> DNA
 <213> Artificial Sequence

5 <220>
 <223> An oligonucleotide

<400> 45
 atattcaggt gcgcatagcg cag 23

10 <210> 46
 <211> 34
 <212> DNA
 <213> Artificial Sequence

15 <220>
 <223> An oligonucleotide

<400> 46
 20 ggacggatat tcagggtcc atagcgaga cttg 34

<210> 47
 <211> 33
 <212> DNA
 25 <213> Artificial Sequence

<220>
 <223> An oligonucleotide

30 <400> 47
 gacggatatt caggtaacca tagcgagac ttg 33

<210> 48
 <211> 30
 35 <212> DNA
 <213> Artificial Sequence

<220>
 <223> An oligonucleotide

40

<400> 48
 ggtgttttca aagtcgcgaa cccgatcgac 30

5 <210> 49
 <211> 26
 <212> DNA
 <213> Artificial Sequence

10 <220>
 <223> An oligonucleotide

<400> 49
 ctgttcagcg aaacttcggg cctgct 26

15 <210> 50
 <211> 26
 <212> DNA
 <213> Artificial Sequence

20 <220>
 <223> An oligonucleotide

<400> 50
 ctgttcagcg aaggttcggg cctgct 26

25 <210> 51
 <211> 26
 <212> DNA
 <213> Artificial Sequence

30 <220>
 <223> An oligonucleotide

<400> 51
 35 ctgttcagcg aagcttcggg cctgct 26

<210> 52
 <211> 22
 <212> DNA
 40 <213> Artificial Sequence

<220>

<223> An oligonucleotide

<400> 52

5 ttcagcgaac gctcgggcct gc

22

<210> 53

<211> 30

<212> DNA

10 <213> Artificial Sequence

<220>

<223> An oligonucleotide

15 <400> 53

ggcctgctgg cttcgcggaa agcgacgaca

30

<210> 54

<211> 21

20 <212> DNA

<213> Artificial Sequence

<220>

<223> An oligonucleotide

25

<400> 54

gaaagcgacg ccaatgacaa t

21

<210> 55

30 <211> 30

<212> DNA

<213> Artificial Sequence

<220>

35 <223> An oligonucleotide

<400> 55

acgacaatga caattgggaa acagcttcgc

30

40 <210> 56

<211> 2265

<212> DNA

<213> Artificial Sequence

5 <220>

<223> A modified DNA sequence

<400> 56

| | | | | | | | |
|----|-------------|-------------|------------|------------|------------|-------------|------|
| | gagggtagag | aaatcgaatg | ccccttgcac | caaggctcgg | ttgacgtttg | cacaggcaaaa | 60 |
| 10 | gccctgtg | cacccgtgac | acagaacatc | aaaacatata | cagtcaagat | tgagaacctg | 120 |
| | cgcgtaatga | ttgatttgag | ctaagaattt | taacaggagg | caccccgggc | cctagagcgt | 180 |
| | aatcaccccc | attccatctt | ttttaggtga | aaacatgaat | tacaataata | aaatcttggg | 240 |
| | aagtgaatct | ggctctgagcc | aaaagcacct | gattcatggc | gatgaagaac | ttttccaaca | 300 |
| | tgaactgaaa | accatttttg | cgcggaactg | gctttttctc | actcatgata | gcctgattcc | 360 |
| 15 | tgcccccg | gactatgtta | ccgcaaaaat | ggggattgac | gaggtcatcg | tctcccgga | 420 |
| | gaacgacgg | tgcattcgtg | cttttctgaa | cgtttgccgg | catcgtggca | agacgctggg | 480 |
| | gagcgtggaa | gccggcaatg | ccaaagggtt | tgtttgccgc | tatcacggct | ggggcttcgg | 540 |
| | ctccaacgg | gaactgcaga | gcgttccatt | tgaaaaagat | ctgtacggcg | agtcgctcaa | 600 |
| | taaaaaatgt | ctggggttga | aagaagtcgc | tcgcgtggag | agcttccatg | gcttcatcta | 660 |
| 20 | cggttgcttc | gaccaggagg | cccctcctct | tatggactat | ctgggtgacg | ctgcttggtta | 720 |
| | cctggaacct | atgttcaagc | attccggcgg | tttagaactg | gtcggtcctc | caggcaagg | 780 |
| | tgtgatcaag | gccaactgga | aggcacccgc | ggaaaacttt | gtgggagatg | cataccacgt | 840 |
| | gggttgga | cacgcgtctt | cgcttcgctc | gggggagctt | atcttctcgt | cgctcgctgg | 900 |
| | caatgcggcg | ctaccacctg | aaggcgcagg | cttgcaaatg | acctccaaat | acggcagcgg | 960 |
| 25 | catgggtgtg | ttgtgggacg | gatattcagg | tgtgcatagc | gcagacttgg | ttccggaatt | 1020 |
| | gatggcattc | ggaggcgcaa | agcaggaaag | gctgaacaaa | gaaattggcg | atgttcgcgc | 1080 |
| | tcggatttat | cgcagccacc | tcaactgcac | cgttttcccg | aacaacagca | tgctgacctg | 1140 |
| | ctcgggtgtt | ttcaaagtat | ggaacccgat | cgacgcaaac | accaccgagg | tctggacctta | 1200 |
| | cgccattgtc | gaaaaagaca | tgcttgagga | tctcaagcgc | cgcttgggcg | actctgttca | 1260 |
| 30 | gcgaacgtgg | gggcctgctg | gcttctggga | aagcgacgac | aatgacaata | tggaacagc | 1320 |
| | ttcgcaaaac | ggcaagaaat | atcaatcaag | agatagtgat | ctgctttcaa | accttggttt | 1380 |
| | cgggtgaggac | gtatacggcg | acgcggtcta | tccaggcgct | gtcggcaaat | cggcgatcgg | 1440 |
| | cgagaccagt | tatcgtgggt | tctaccgggc | ttaccaggca | cacgtcagca | gctccaactg | 1500 |
| | ggctgagttc | gagcatgcct | ctagtacttg | gcatactgaa | cttacgaaga | ctactgatcg | 1560 |
| 35 | ctaacagacg | agtcgacct | gatgatcaat | attcaagaag | acaagctgg | ttccgcccac | 1620 |
| | gacgccgaag | agattcttcg | tttcttcaat | tgccacgact | ctgctttgca | acaagaagcc | 1680 |
| | actacgctgc | tgaccaggga | agcgcatctg | ttggacattc | aggcttaccg | tgcttggtta | 1740 |
| | gagcactgcg | tggggtcaga | ggtgcaatat | caggtcattt | cacgcgaact | gcgcgcagct | 1800 |
| | tcagagcgtc | gttataagct | caatgaagcc | atgaacgttt | acaacgaaaa | ttttcagcaa | 1860 |
| 40 | ctgaaagtcc | gagttgagca | tcaactggat | ccgcaaaact | ggggcaacag | cccgaagctg | 1920 |

cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaataa caaagagcta 1980
 cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggatgatgtc 2040
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggtc 2100
 cagcgattcg tcgattaccc agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
 5 tgattcagtg accattttta caaatgggtca ctgcaaccgc ggatcaccatt aatcaaaggg 2220
 aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg 2265

<210> 57

<211> 2265

10 <212> DNA

<213> Artificial Sequence

<220>

<223> A modified DNA sequence

15

<400> 57

gagggtagag aaatcgaatg ccccttgcat caaggctcgt ttgacgtttg cacaggcaaa 60
 gccctgtgag caccctgtgac acagaacatc aaaacatata cagtcaagat tgagaacctg 120
 cgcgtaatat ttgatttgag ctaagaattt taacaggagg caccctgggc cctagagcgt 180
 20 aatcaccccc attccatctt ttttaggtga aaacatgaat tacaataata aaatcttggt 240
 aagtgaatct ggtctgagcc aaaagcacct gattcatggc gatgaagaac ttttccaaca 300
 tgaactgaaa accatttttg cgcggaactg gctttttctc actcatgata gcctgattcc 360
 tgccccgggc gactatgtta ccgcaaaaat ggggattgac gaggtcatcg tctccccgca 420
 gaacgacggt tcgattcgtg cttttctgaa cggttgccgg catcgaggca agacgctggt 480
 25 gagcgtggaa gccggaatg ccaaagggtt tgtttgcagc tatcacggct ggggcttcgg 540
 ctccaacggt gaactgcaga gcgttcattt tgaaaaagat ctgtacggcg agtcgctcaa 600
 taaaaaatgt ctgggggtga aagaagtcgc tcgcgtggag agcttccatg gttcatcta 660
 cgggtgcttc gaccaggagg cccctcctct tatggactat ctgggtgacg ctgcttggtg 720
 cctggaacct atgttcaagc attccggcgg tttagaactg gtcggtcctc caggcaaggt 780
 30 tgtgatcaag gccaactgga aggcacccgc ggaaaacttt gtgggagatg cataccacgt 840
 ggggttgacg cacgcgtctt cgcttcgctc gggggagtct atcttctcgt cgctcgctgg 900
 caatgcggcg ctaccacctg aaggcgcagg cttgcaaatg acctccaaat acggcagcgg 960
 catgggtgtg ttgtgggacg gatattcagg tgtgcatagc gcagacttgg ttccggaatt 1020
 gatggcattc ggaggcgcaa agcaggaaaag gctgaacaaa gaaattggcg atgttcgcgc 1080
 35 tcggatttat cgcagccacc tcaactgcac cgttttcccg aacaacagca tgctgacctg 1140
 ctccgggtgtt ttcaaagtat ggaacccgat cgacgcaaac accaccgagg tctggacctg 1200
 cgccattgtc gaaaaagaca tgcttgagga tctcaagcgc cgcttgggcg actctgttca 1260
 gcgaacgggc gggcctgctg gcttctggga aagcgacgac aatgacaata tggaaacagc 1320
 ttcgcaaaac ggcaagaaat atcaatcaag agatagtgat ctgctttcaa accttggttt 1380
 40 cggtgaggac gtatacggcg acgcggtcta tccaggcgtc gtcggcaaat cggcgatcgg 1440

cgagaccagt tatcgtgggt tctaccgggc ttaccaggca cacgtcagca gctccaactg 1500
 ggctgagttc gagcatgcct ctactacttg gcatactgaa cttacgaaga ctactgatcg 1560
 ctaacagacg agtcgaccat gatgatcaat attcaagaag acaagctggt ttccgcccac 1620
 gacgccgaag agattcttcg tttcttcaat tgccacgact ctgctttgca acaagaagcc 1680
 5 actacgctgc tgacccagga agcgcatctt ttggacattc aggcttaccg tgcttggtta 1740
 gagcactgcg tggggtcaga ggtgcaatat caggtcattt cacgcgaact gcgcgcagct 1800
 tcagagcgtc gttataagct caatgaagcc atgaacgttt acaacgaaaa ttttcagcaa 1860
 ctgaaagttc gagttgagca tcaactggat ccgcaaaact ggggcaacag cccgaagctg 1920
 cgctttactc gctttatcac caacgtccag gccgcaatgg acgtaaatga caaagagcta 1980
 10 cttcacatcc gctccaacgt cattctgcac cgggcacgac gtggcaatca ggtcgatgct 2040
 ttctacgccg cccgggaaga taaatggaaa cgtggcgaag gtggagtacg aaaattggct 2100
 cagcgattcg tcgattaccg agagcgcata cttcagacgc acaatctgat ggtctttctg 2160
 tgattcagtg accattttta caaatggtca ctgcaaccgc ggtcaccatt aatcaaaggg 2220
 aatgtacgtg tatgggcaat caacaagtcg tttcgataac cggtg 2265

15

<210> 58
 <211> 449
 <212> PRT
 <213> Artificial Sequence

20

<220>
 <223> A polypeptide encoded by SEQ ID NO:56

<400> 58

25 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1 5 10 15
 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
 20 25 30
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
 30 35 40 45
 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
 50 55 60
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
 65 70 75 80
 35 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
 85 90 95
 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
 100 105 110
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
 40 115 120 125

Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 130 135 140
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
 145 150 155 160
 5 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
 165 170 175
 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
 180 185 190
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
 10 195 200 205
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
 210 215 220
 Ser Ser Leu Ala Gly Asn Ala Ala Leu Pro Pro Glu Gly Ala Gly Leu
 225 230 235 240
 15 Gln Met Thr Ser Lys Tyr Gly Ser Gly Met Gly Val Leu Trp Asp Gly
 245 250 255
 Tyr Ser Gly Val His Ser Ala Asp Leu Val Pro Glu Leu Met Ala Phe
 260 265 270
 Gly Gly Ala Lys Gln Glu Arg Leu Asn Lys Glu Ile Gly Asp Val Arg
 20 275 280 285
 Ala Arg Ile Tyr Arg Ser His Leu Asn Cys Thr Val Phe Pro Asn Asn
 290 295 300
 Ser Met Leu Thr Cys Ser Gly Val Phe Lys Val Trp Asn Pro Ile Asp
 305 310 315 320
 25 Ala Asn Thr Thr Glu Val Trp Thr Tyr Ala Ile Val Glu Lys Asp Met
 325 330 335
 Pro Glu Asp Leu Lys Arg Arg Leu Ala Asp Ser Val Gln Arg Thr Trp
 340 345 350
 Gly Pro Ala Gly Phe Trp Glu Ser Asp Asp Asn Asp Asn Met Glu Thr
 30 355 360 365
 Ala Ser Gln Asn Gly Lys Lys Tyr Gln Ser Arg Asp Ser Asp Leu Leu
 370 375 380
 Ser Asn Leu Gly Phe Gly Glu Asp Val Tyr Gly Asp Ala Val Tyr Pro
 385 390 395 400
 35 Gly Val Val Gly Lys Ser Ala Ile Gly Glu Thr Ser Tyr Arg Gly Phe
 405 410 415
 Tyr Arg Ala Tyr Gln Ala His Val Ser Ser Ser Asn Trp Ala Glu Phe
 420 425 430
 Glu His Ala Ser Ser Thr Trp His Thr Glu Leu Thr Lys Thr Thr Asp
 40 435 440 445

Arg

<210> 59
 5 <211> 449
 <212> PRT
 <213> Artificial Sequence

 <220>
 10 <223> A polypeptide encoded by SEQ ID NO:57

<400> 59
 Met Asn Tyr Asn Asn Lys Ile Leu Val Ser Glu Ser Gly Leu Ser Gln
 1 5 10 15
 15 Lys His Leu Ile His Gly Asp Glu Glu Leu Phe Gln His Glu Leu Lys
 20 25 30
 Thr Ile Phe Ala Arg Asn Trp Leu Phe Leu Thr His Asp Ser Leu Ile
 35 40 45
 Pro Ala Pro Gly Asp Tyr Val Thr Ala Lys Met Gly Ile Asp Glu Val
 20 50 55 60
 Ile Val Ser Arg Gln Asn Asp Gly Ser Ile Arg Ala Phe Leu Asn Val
 65 70 75 80
 Cys Arg His Arg Gly Lys Thr Leu Val Ser Val Glu Ala Gly Asn Ala
 85 90 95
 25 Lys Gly Phe Val Cys Ser Tyr His Gly Trp Gly Phe Gly Ser Asn Gly
 100 105 110
 Glu Leu Gln Ser Val Pro Phe Glu Lys Asp Leu Tyr Gly Glu Ser Leu
 115 120 125
 Asn Lys Lys Cys Leu Gly Leu Lys Glu Val Ala Arg Val Glu Ser Phe
 30 130 135 140
 His Gly Phe Ile Tyr Gly Cys Phe Asp Gln Glu Ala Pro Pro Leu Met
 145 150 155 160
 Asp Tyr Leu Gly Asp Ala Ala Trp Tyr Leu Glu Pro Met Phe Lys His
 165 170 175
 35 Ser Gly Gly Leu Glu Leu Val Gly Pro Pro Gly Lys Val Val Ile Lys
 180 185 190
 Ala Asn Trp Lys Ala Pro Ala Glu Asn Phe Val Gly Asp Ala Tyr His
 195 200 205
 Val Gly Trp Thr His Ala Ser Ser Leu Arg Ser Gly Glu Ser Ile Phe
 40 210 215 220

Arg
30

<211> 26

35 <213> Artificial Sequence

<223> An oligonucleotide

40 <400> 60

gttcagcgaa cgggcgggcc tgctgg

26

<210> 61

<211> 26

5 <212> DNA

<213> Artificial Sequence

<220>

<223> An oligonucleotide

10

<400> 61

gttcagcgaa cggccgggcc tgctgg

26

<210> 62

15 <211> 26

<212> DNA

<213> Artificial Sequence

<220>

20 <223> An oligonucleotide

<400> 62

gttcagcgaa cgaccgggcc tgctgg

26

25 <210> 63

<211> 26

<212> DNA

<213> Artificial Sequence

30 <220>

<223> An oligonucleotide

<400> 63

gttcagcgaa cgatcgggcc tgctgg

26

35

<210> 64

<211> 26

<212> DNA

<213> Artificial Sequence

40

<223> An oligonucleotide

5 gttcagcgaa cgtggggggcc tgctgg

[illegible]

<211> 25

10 <213> Artificial Sequence

<223> An oligonucleotide

ttcagcgaac gtacgggcct gctgg

25